Записки

по

Теория на екстремалните задачи

Борислав Драганов

София, 2023

Това са записки на лекциите по теория на екстремалните задачи, които чета във Факултета по математика и информатика на Софийския университет. В курса се разглеждат задачи за екстремум на изображения и оператори при наличието на различни условия. Те обобщават познатите задачи за намиране на безусловен или условен екстремум на функции на една или няколко поменливи, познати от курсовете по диференциално и интегрално смятане.

Лекциите се основават почти изключително на книгата на **Йофе и Тихомиров "Теория экстремаьлных задач"**. Курсът е включен в магистърската програма "Оптимизация" и е замислен така, че да се допълва с други курсове в нея — курсове по вариационно смятане, оптимално управление и оптимиране. Поради това в настоящите лекции се обръща повече внимание на чисто теоретичните аспекти на дисциплината и е отделено съвсем малко внимание на иначе изключително разнообразните и важни приложения на този дял от математическия анализ.

Важни предпоставки за изучаване на курса са познаването и владеенето на редица основни понятия от функционалния анализ. На това се обръща внимание в първата лекция. Втората лекция е посветена на диференциалното смятане за оператори (или изображения). То играе основна роля в изграждането на теорията на гладките екстремални задачи. След полагането на тази основа, се преминава към основните екстремални задачи за гладки функции — лекции 3, 4 и 5. Те са предмет и на вариационното смятане. По-нататък преминаваме към екстремални задачи, в които предположението за гладкост е заместено с това за изпъкналост. За тази цел в лекция 6 първо се припомнят нужните сведения от изпъкналия анализ, а след това, в лекция 7, се излага основният резултат за този тип екстремални задачи. В лекция 8 пък се разглеждат задачи, които комбинират в себе си характеристики на споменатите по-горе две. Лекции 9 и 10 съдържат основния резултат в теория на екстремалните задачи — Принципа за максимума. Той дава необходимо условие за екстремум в задачи, които се изучават в оптималното управление. В лекция 11 се формулира едно усложнение в условията на такива задачи. Дотук се извеждат само необходими условия за екстремум. В лекция 12 се излагат няколко основни достатъчни условия. Последната 13-та лекция съдържа решения на конкретни екстремални задачи. Решаването на задачи е много важна част от вникването в теорията, затова всички лекции завършват със задачи (изключение прави Лекция 10, която е посветена на доказателството на Принципа за максимума). Броят им нарочно не е голям — надявам се това да бъде допълнителен стимул да им обърнете внимание.

Благодарен съм на студентите, посетили курса, за корекции, отзиви и предложения за подобрения на изложението. Неминуемо в записките все още са останали грешки и неточности. Ще бъда много благодарен на всеки, който ми съобщи за такива. Може да ми пишете на `bdraganov@fmi.uni-sofia.bg`. Разбира се, много ценни ще бъдат и предложенията за подобрения.

Съдържание

1 Представяне на линейни функционали. Теореми за отделимост

Ще направим кратък преглед на основните понятия и теореми от функционалния анализ, които ще използваме по-късно при представянето на теорията на екстремалните задачи. Ще предполагаме, че сме запознати с това какво е линейно топологично пространство, норма, пълно пространство, компакт, линеен оператор и линеен функционал. Винаги разглеждаме линейни пространства над полето на реалните числа.[1]

1.1 Представяне на ограничени линейни функционали

Нека X е линейно топологично пространство, а ℓ е линеен фунцкционал върху X. За удобство обикновено действието на ℓ върху $x \in X$ ще означаваме чрез $\langle \ell, x \rangle$ вместо чрез $\ell(x)$ или $\ell\, x$.

Нека $a, b \in \mathbb{R}^n$, като $a := (a_1, a_2, \dots, a_n)$ и $b := (b_1, b_2, \dots, b_n)$. Стандартното скаларно произведение на a и b означаваме с $a \cdot b$, т.е.

$$a \cdot b := \sum_{i=1}^{n} a_i b_i.$$

Известно е, че линейните функционали ℓ върху $\mathbb{R}^n$ се представят във вида

$$\langle \ell, a \rangle = l \cdot a, \quad a \in \mathbb{R}^n, \tag{1.1}$$

за някой $l \in \mathbb{R}^n$.

Нека $[a, b] \subset \mathbb{R}$. Според теорема на Рис, всеки ограничен линеен функционал ℓ върху $C[a, b]$ се представя по единствен начин във вида

$$\langle \ell, x \rangle = \int_a^b x(t)\, d\mu, \quad x \in C[a, b],$$

с някоя знакопроменлива крайна регулярна борелова мярка μ върху $[a, b]$, или еквивалентно

$$\langle \ell, x \rangle = \int_a^b x(t)\, d\mu(t) \tag{1.2}$$

с някоя функция с ограничена вариация $\mu(t)$ върху $[a, b]$, която е непрекъсната отляво в (a, b) и $\mu(a) = 0$. Една крайна борелова мярка μ върху интервал се нарича *регулярна*, ако за всяко борелово множество S и всяко $\varepsilon > 0$ съществуват затворено множество $C \subseteq S$ и отворено множество

[1] Тези, които не са запознати с понятията топология и топологично пространство, навсякъде биха могли да си мислят, че става дума за нормирани линейни пространства.

$O \supseteq S$ такива, че $\mu(O \backslash C) < \varepsilon$. Условието за регулярност е еквивалентно на

$$\mu(S) = \inf\{\mu(O) : O \supseteq S,\ O \text{ е отворено}\},$$
$$\mu(S) = \sup\{\mu(C) : C \subseteq S,\ C \text{ е затворено}\}$$

за всяко борелово множество S.

Нека още припомним, че *носител* на дадена неотрицателна борелова мярка върху интервал наричаме най-широкото затворено множество със свойството всяка отворена околност на всяка точка от него да има положителна мярка.

Според една друга теорема на Рис ограничените линейни функционали над $L[a, b]$ се представят във вида

$$\langle \ell, x\rangle = \int_a^b s(t)\, x(t)\, dt, \quad x \in L[a, b], \tag{1.3}$$

с някоя функция $s \in L_\infty[a, b]$.

Да изведем сега обобщението на тези теореми за векторно-значни функции. Пространството от непрекъснатите функции, дефинирани в интервала $[a, b]$ със стойности в $\mathbb{R}^n$, ще означаваме с $C([a, b], \mathbb{R}^n)$, а на сумириемите (по Лебег) — с $L([a, b], \mathbb{R}^n)$.

Дефинираме римановия, съответно лебеговия, интеграл на векторно-значната функция $x(t) := (x_1(t), x_2(t), \dots, x_n(t))$ от $C([a, b], \mathbb{R}^n)$, съответно $L([a, b], \mathbb{R}^n)$, като полагаме

$$\int_a^b x(t)\, dt := \left(\int_a^b x_1(t)\, dt, \dots, \int_a^b x_n(t)\, dt \right).$$

От (1.1) и (1.2) следва, че всеки ограничен линеен функционал ℓ върху пространството $C([a, b], \mathbb{R}^n)$ се представя във вида

$$\langle \ell, x\rangle = \sum_{i=1}^n \int_a^b x_i(t)\, d\mu_i(t), \quad x := (x_1, x_2, \dots, x_n) \in C([a, b], \mathbb{R}^n), \tag{1.4}$$

където $\mu_i(t)$ са функции с ограничена вариация върху $[a, b]$. Ако въведем означението,

$$d\mu(t) := (d\mu_1(t), d\mu_2(t), \dots, d\mu_n(t)),$$

горната формула може да се запише във вида

$$\langle \ell, x\rangle = \int_a^b x(t) \cdot d\mu(t). \tag{1.5}$$

Аналогично от (1.1) и (1.3) следва, че всеки ограничен линеен функционал ℓ върху пространството $L([a, b], \mathbb{R}^n)$ се представя във вида

$$\langle \ell, x\rangle = \int_a^b s(t) \cdot x(t)\, dt, \quad x := (x_1, x_2, \dots, x_n) \in L([a, b], \mathbb{R}^n), \tag{1.6}$$

където $s \in L_\infty([a, b], \mathbb{R}^n)$.

1.2 Теореми за отделимост

Нека X е линейно топологично пространство. Съвкупността от всички непрекъснати линейни функционали върху X ще означаваме с X^*. Ако X е нормирано, с $\|x^*\|$ ще означаваме стандартната (операторна) норма на $x^* \in X^*$, т.е.

$$\|x^*\| := \sup_{\substack{x \in X \\ x \neq 0}} \frac{|\langle x^*, x\rangle|}{\|x\|_X}.$$

Да припомним, че за линейните функционали, по-общо оператори, действащи между нормирани пространства, свойствата непрекъснатост и ограниченост са еквивалентни.

Преминаваме към формулировките на теоремата на Хан-Банах за отделимост и основните ѝ следствия.

Теорема 1.1 (Хан-Банах)**.** *Нека X е линейно топологично пространство, $A \subseteq X$ е отворено изпъкнало множество и $L \subseteq X$ е линейно подпространство, което няма общи точки с A. Тогава съществува $x^* \in X^*$ такова, че*

$$\begin{aligned} \langle x^*, x\rangle &> 0 \quad x \in A, \\ \langle x^*, x\rangle &= 0 \quad x \in L. \end{aligned}$$

Ще приведем също и няколко важни следствия на тази теорема.

Следствие 1.2. *Нека X е хаусдорфово локално изпъкнало топологично пространство. Тогава за всяко $x \in X$, $x \neq 0$, съществува $x^* \in X^*$ такъв, че $\langle x^*, x\rangle \neq 0$.*

Да припомним, че, ако L е линейно подпространство на линейното топологично пространство X, то множеството

$$L^\perp := \{x^* \in X^* : \langle x^*, x\rangle = 0 \quad \forall x \in L\}$$

се нарича *анулатор* на L.

Следствие 1.3. *Анулаторът на всяко същинско линейно подпространство на хаусдорфово локално изпъкнало линейно топологично пространство съдържа ненулев елемент.*

Следствие 1.4. *Нека X е нормирано линейно пространство. Тогава за всяко $x \in X$ съществува $x^* \in X^*$ такъв, че $\langle x^*, x\rangle = \|x\|_X$ и $\|x^*\| = 1$.*

Теоремата на Хан-Банах може да се формулира и в термините на отделимост. Казваме, че непрекъснатият линеен функционал x^* над X *разделя* множествата $A, B \subseteq X$, ако $\langle x^*, a\rangle \leq \langle x^*, b\rangle$ за всеки $a \in A$ и $b \in B$.

Множеството от вътрешните точки на множеството A ще означаваме с $\operatorname{int} A$.

Теорема 1.5 (Хан-Банах). *Нека A и B са непразни, непресичащи се, изпъкнали подмножества на линейно топологично пространство X, като $\operatorname{int} A \neq \emptyset$. Тогава съществува ненулев $x^* \in X^*$, който разделя множествата A и B, т.е. такъв, че*

$$\langle x^*, a\rangle \leq \langle x^*, b\rangle \quad \forall a \in A \quad \forall b \in B.$$

В крайномерни пространства разполагаме с по-силно твърдение — отпада предположението, че едно от множествата има непразна вътрешност.

Теорема 1.6. *Нека A и B са непразни, непресичащи се, изпъкнали подмножества на $\mathbb{R}^n$. Тогава съществува ненулев ограничен линеен функционал x^* над $\mathbb{R}^n$, който разделя множествата A и B.*

Ще докажем едно важно свойство на анулатора. Първо да припомним, че ако X и Y са линейни топологични пространства и $\Lambda : X \to Y$ е непрекъснат линеен оператор, то неговият *спрегнат* оператор $\Lambda^* : Y^* \to X^*$ се дефинира посредством съотношението

$$\langle \Lambda^* y^*, x\rangle = \langle y^*, \Lambda x\rangle \quad y^* \in Y^*,\ x \in X.$$

За изображение $F : X \to Y$ с $\operatorname{Im} F$ ще означаваме неговата област от стойности, т.е. $\operatorname{Im} F := F(X)$, а с $\operatorname{Ker} F$ — неговото ядро, т.е. $\operatorname{Ker} F := \{x \in X : F(x) = 0\}$.

Лема 1.7 (Лема за анулатора). *Нека X и Y са линейни топологични пространства и $\Lambda : X \to Y$ е непрекъснат линеен оператор такъв, че $\operatorname{Im} \Lambda = Y$. Тогава $(\operatorname{Ker} \Lambda)^\perp = \operatorname{Im} \Lambda^*$.*

Доказателство. Ако $x^* \in \operatorname{Im} \Lambda^*$, т.е. съществува $y^* \in Y^*$ такова, че $x^* = \Lambda^* y^*$, то за всяко $x \in \operatorname{Ker} \Lambda$ имаме

$$\langle x^*, x\rangle = \langle \Lambda^* y^*, x\rangle = \langle y^*, \Lambda x\rangle = 0.$$

Следователно $x^* \in (\operatorname{Ker} \Lambda)^\perp$. Така установихме включването $\operatorname{Im} \Lambda^* \subseteq (\operatorname{Ker} \Lambda)^\perp$.

Обратно, нека $x^* \in (\operatorname{Ker} \Lambda)^\perp$. Трябва да покажем, че съществува $y^* \in Y^*$ такова, че $x^* = \Lambda^* y^*$. За да се ориентираме как да дефинираме такова y^*, можем да забележим, че ако то съществува, ще имаме $\langle y^*, \Lambda x\rangle = \langle \Lambda^* y^*, x\rangle = \langle x^*, x\rangle$. Това ни подсказва да дефинираме $y^* \in Y^*$, полагайки $\langle y^*, y\rangle := \langle x^*, x\rangle$, където $x \in X$е такова, че $y = \Lambda x$. Такова x съществува, защото $\operatorname{Im} \Lambda = Y$.

Така, за да завършим доказателството, остава да се убедим, че горната дефиниция е коректна. Това е така, защото, ако $\Lambda x_1 = \Lambda x_2 = y$, то $x_1 - x_2 \in \operatorname{Ker}\Lambda$. Следователно $\langle x^*, x_1 - x_2\rangle = 0$, защото $x^* \in (\operatorname{Ker} \Lambda)^\perp$. Оттук следва, че $\langle x^*, x_1\rangle = \langle x^*, x_2\rangle$. □

Задачи

1. Докажете представянето (1.1).

2. Изведете от Теорема 1.1 формулираните след нея три следствия.

2 Диференциално смятане за изображения

2.1 Дефиниции

Ще дадем дефинициите на трите основни обобщения на понятието производна. Те са: производна по посока, производна на Гато́ (или още слаба производна) и производна на Фреше́ (или още силна производна).

Дефиниция 2.1. (производна по посока) Нека X и Y са линейни топологични пространства. Нека $U \subseteq X$ е отворено, $x_0 \in U$, $h \in X$ и $F : U \to Y$. Ако границата

$$F'(x_0, h) := \lim_{t \to 0+} \frac{F(x_0 + th) - F(x_0)}{t}$$

съществува,[1] то казваме, че изображението F е диференцируемо в точката x_0 по посоката h, а $F'(x_0, h)$ наричаме *производна* на F в точката x_0 *по посоката* h.

Според дефиницията, производната по посока $F'(x_0, h)$, стига да съществува, е елемент на Y. Нека отбележим, че щом U е отворено, то x_0 е негова вътрешна точка и тогава $x_0 + th \in U$ за всяко положително достатъчно малко t и $F(x_0 + th)$ е добре дефинирано.

Бележка 2.2. Нека $F : U \to \mathbb{R}$, т.е. $Y = \mathbb{R}$. Ако положим $\varphi(t) := F(x_0 + th)$ за $t \geq 0$, то имаме, че изображението F е диференцируемо в точката x_0 по посоката h тогава и само тогава, когато $\varphi(t)$ е диференцируемо в точката 0 отдясно (в познатия смисъл на реално-значни функции на една реална променлива), при това $F'(x_0, h) = \varphi'(0+)$.

Дефиниция 2.3. (първа вариация) В условията на предходната дефиниция, ако $F'(x_0, h)$ съществува за всяко $h \in X$, казваме, че F има *първа вариация* в точката x_0 и я означаваме с $\delta F(x_0)$, като полагаме

$$\delta F(x_0)(h) := F'(x_0, h), \quad h \in X.$$

Да отбележим, че първата вариация (ако съществува) е изображение (оператор) от X в Y.

Дефиниция 2.4. (производна на Гато́) В означнията на предходните дефиниции, ако $\delta F(x_0)$ е непрекъснат (еквивалентно, ограничен) линеен оператор от X в Y, то казваме, че изображението F е диференцируемо по Гато́ (или слабо диференцируемо) в точката x_0 и полагаме

$$F'_G(x_0)\, h := \delta F(x_0)(h), \quad h \in X.$$

Непрекъснатият линеен оператор $F'_G(x_0) : X \to Y$ се нарича *производна на Гато́* на F в точката x_0.

[1] В смисъла на топологията на Y или, в случай че Y е нормирано, в неговата норма.

Тук и по-нататък следваме конвенцията да означаваме аргумента на линейните оператори (в частност, функционали) без скоби, т.е. ако Λ е линеен оператор, пишем Λx вместо $\Lambda(x)$; освен, разбира се, когато прилагаме Λ към по-сложен израз, например, сума.

Стигнахме до последното понятие. То е най-тясно свързано с познатата производна на функция.

Дефиниция 2.5. (производна на Фреше́) Нека X и Y са нормирани линейни пространства с норми съответно $\| \circ \|_X$ и $\| \circ \|_Y$. Нека $U \subseteq X$ е отворено и $F : U \to Y$. Казваме, че изображението F е диференцируемо по Фреше́ (или силно диференцируемо) в точката $x_0 \in U$, ако съществува непрекъснат линеен оператор $\Lambda : X \to Y$ такъв, че

$$\frac{\|F(x_0 + h) - F(x_0) - \Lambda h\|_Y}{\|h\|_X} \to 0 \quad \text{при} \quad \|h\|_X \to 0.$$

В такъв случай, полагаме $F'(x_0) := \Lambda$ и го наричаме *производна на Фреше́* на изображението F в точката x_0.

Непосредствено от дефиницията горе следва, че ако производната на Фреше́ на даден оператор в дадена вътрешна точка съществува, то тя е единствена.

Също така се вижда, че стига да съществува, производната на Фреше́ $F'(x_0)$ на $F(x)$ в точката x_0 е такъв ограничен линеен оператор от X в Y, че е в сила съотношението

$$F(x_0 + h) = F(x_0) + F'(x_0)h + r(h), \tag{2.1}$$

където $r : X \to Y$ е такова, че $\|r(h)\|_Y = o(\|h\|_X)$. Накратко, понякога записваме (2.1) по следния начин

$$F(x_0 + h) = F(x_0) + F'(x_0)h + o(\|h\|_X).$$

Забележете, че в случая на изображения на няколко променливи със стойности в крайномерно евклидово пространство горната дефиниция съвпада с познатата за пълна производна на изображение (понякога наричана матрична производна).

Нека X и Y са нормирани линейни пространства. Да припомним, че едно изображение $B(x_1, x_2)$, дефинирано в $X \times X$ със стойности в Y, се нарича:

- *билинейно*, ако то е линейно както по x_1, така и по x_2;
- *ограничено*, ако съществува положително реално число c такова, че $\|B(x_1, x_2)\|_Y \le c\,\|x_1\|_X\,\|x_2\|_X$ за всеки $x_1, x_2 \in X$;
- *симетрично*, ако $B(x_1, x_2) = B(x_2, x_1)$ за всеки $x_1, x_2 \in X$.

Дефиниция 2.6. Нека X и Y са нормирани линейни пространства, $U \subseteq X$ е отворено и $F : U \to Y$. Казваме, че изображението $F(x)$ притежава втора производна на Фреше́ в точката $x_0 \in U$, ако е диференцируемо по Фреше́ в x_0 и съществува симетричен ограничен билинеен оператор $B : X \times X \to Y$ такъв, че

$$F(x_0 + h) = F(x_0) + F'(x_0)h + \frac{1}{2}B(h, h) + r(h),$$

където $\|r(h)\|_Y = o(\|h\|_X^2)$. Операторът B се означава с $F''(x_0)$ и се нарича *втора производна на Фреше́* на $F(x)$ в x_0.

Нека X и Y са нормирани линейни пространства, $U \subseteq X$ е отворено и $F : U \to Y$. Нека производната на Фреше́ на $F(x)$ съществува в околност $V \subseteq U$ на точката x_0. Тогава $F' : V \to L(X, Y)$, където с $L(X, Y)$ сме означили пространството на непрекъснатите линейни оператори от X в Y, нормирано с обичайната операторна норма. Може да се покаже, че ако $F'(x)$, разгледано като изображение, зависещо от x, е диференцируемо в смисъл на Фреше́ в x_0, то неговата производна на Фреше́ в тази точка е $F''(x_0)$, т.е. $(F')'(x_0) = F''(x_0)$.

Сравнявайки различните видове производни, които въведохме, можем да кажем, че започваме с производна по посока $F'(x_0, h)$. Ако тя съществува по всяка посока h, я наричаме първа вариация на F в x_0. По-нататък, ако $F'(x_0, h)$, освен че съществува за всяко h, представлява непрекъснат линеен оператор по h, то наричаме този оператор призводна на Гато́ на F в точката x_0. Накрая, както можем да се убедим, ако сходимимостта в дефиницията на производна на Гато́ е равномерна по единичното кълбо в X, то F е диференцируемо по Фреше́ в x_0 и производните му на Гато́ и Фреше́ съвпадат.

2.2 Основни свойства

Преминаваме към основните свойства на въведените производни. От (2.1) веднага следва, че диференцируемост по Фреше́ влече непрекъснатост.

Твърдение 2.7. *Ако изображение е диференцируемо по Фреше́ в дадена точка, то то е непрекъснато в нея.*

Друго свойство, което веднага следва от дефиницията за производна на Фреше́, е че всяко непрекъснато линейно изображение (линеен оператор) $\Lambda : X \to Y$ е дифернцируемо и

$$\Lambda'(x) = \Lambda, \quad x \in X,$$

т.е.

$$\Lambda'(x)h = \Lambda h, \quad x, h \in X.$$

Твърдение 2.8. *Ако изображението $F(x)$ е диференцируемо по Фреше́ в точката x_0, то то е диференцируемо и по Гато́ в нея, при това $F'_G(x_0) = F'(x_0)$.*

Доказателство. За достатъчно малко $t > 0$ и ненулево $h \in X$, имаме

$$\begin{aligned}&\left\| \frac{F(x_0 + th) - F(x_0)}{t} - F'(x_0)h \right\|_Y \\ &\qquad\qquad = \|h\|_X \frac{\|F(x_0 + th) - F(x_0) - F'(x_0)(th)\|_Y}{\|th\|_X} \to 0, \quad t \to 0,\end{aligned}$$

защото $F(x)$ е диференцируемо по Фреше́ в x_0. □

Непосредствено от дефинициите следва линейността на всеки един от въведените видове диференциране.

Твърдение 2.9. *Ако изображенията $F_1(x)$ и $F_2(x)$ са диференцируеми във вътрешната точка x_0 по посоката h, съответно по Гато́ или по Фреше́, то и изображението $aF_1(x) + bF_2(x)$, където $a, b \in \mathbb{R}$, също е диференцируемо в x_0 по посоката h, съответно по Гато́ или по Фреше́, като*

$$\big(aF_1 + bF_2\big)'(x_0, h) = aF_1'(x_0, h) + bF_2'(x_0, h),$$

съответно

$$\big(aF_1 + bF_2\big)'_G(x_0) = aF'_{1_G}(x_0) + bF'_{2_G}(x_0)$$

или

$$\big(aF_1 + bF_2\big)'(x_0) = aF_1'(x_0) + bF_2'(x_0).$$

Следва правилото за диференциране по Фреше́ на съставно изображение.

Теорема 2.10. *Нека X, Y и Z са нормирани линейни пространства. Нека U и V са отворени подмножества съответна на X и Y. Нека още $F : U \to Y$ и $G : V \to Z$ са изображения, а $x_0 \in U$ е такава, че $F(x_0) \in V$. Тогава, ако $F(x)$ е диференцируемо по Фреше́ в точката x_0 и $G(y)$ е диференцируемо по Фреше́ в точката $F(x_0)$, то съставното изображение $H(x) := G(F(x))$ също е диференцируемо по Фреше́ в точката x_0, при това*

$$H'(x_0) = G'(F(x_0))F'(x_0).$$

Нека изрично обърнем внимание на това, че в горната формула става дума за композиция на линейните оператори $G'(F(x_0))$ и $F'(x_0)$. Разбира се, в крайномерния случай производните на Фреше́ представляват матрици (спомнете си дефиницията на пълна производна на изображения на краен брой реални променливи със стойности в крайномерно евклидово пространство) и тогава тази формула се свежда до умножението на две матрици, в частност, числа, ако имаме реално-значни функции на една реална променлива.

Доказателство на Твърдение 2.10. Да положим $y_0 := F(x_0)$. От дефиницията на производна на Фреше́ непосредствено следва (вж. (2.1))

$$F(x_0 + \xi) = F(x_0) + F'(x_0)\xi + r(\xi)$$

и

$$G(y_0 + \eta) = G(y_0) + G'(y_0)\eta + R(\eta),$$

където $\|r(\xi)\|_Y = o(\|\xi\|_X)$ и $\|R(\eta)\|_Z = o(\|\eta\|_Y)$. Следователно за ξ достатъчно малко по норма имаме

$$\begin{aligned} H(x_0 + \xi) &= G(F(x_0 + \xi)) = G(y_0 + F'(x_0)\xi + r(\xi)) \\ &= G(y_0) + G'(y_0)(F'(x_0)\xi + r(\xi)) + R(F'(x_0)\xi + r(\xi)) \\ &= H(x_0) + G'(y_0)F'(x_0)\xi + G'(y_0)r(\xi) + R(F'(x_0)\xi + r(\xi)). \end{aligned}$$

При последното равенство използваме, че $G'(y_0)$ е линеен оператор. За да завършим доказателството, остава да покажем, че

$$\|G'(y_0)r(\xi) + R(F'(x_0)\xi + r(\xi))\|_Z = o(\|\xi\|_X).$$

Това следва от неравенството

$$\begin{aligned} &\|G'(y_0)r(\xi) + R(F'(x_0)\xi + r(\xi))\|_Z \\ &\qquad\qquad \leq \|G'(y_0)\|_{Y\to Z}\|r(\xi)\|_Y + \|R(F'(x_0)\xi + r(\xi))\|_Z, \end{aligned}$$

ограничеността на оператора $F'(x_0)$ и съотношенията $\|r(\xi)\|_Y = o(\|\xi\|_X)$ и $\|R(\eta)\|_Z = o(\|\eta\|_Y)$. □

Ако изображение се задава посредством координатни изображения, то диференцируемостта му се свежда до тази на координатните изображения. По-точно, нека $Y_1, \ldots, Y_n$ са нормирани линейни пространства. Тогава линейното пространство $Y := Y_1 \times \cdots \times Y_n$ е нормирано с норма $\|y\|_Y := \max_{i=1,\ldots,n} \|y_i\|_{Y_i}$, където $y := (y_1, \ldots, y_n)$. В сила е следното елементарно свойство.

Твърдение 2.11. *Нека X и $Y_1, \ldots, Y_n$ са нормирани линейни пространства, $U \subseteq X$ е отворено, $F_i : U \to Y_i$, $i = 1, \ldots, n$, и $F : U \to Y$ е дефинирано посредством $F(x) := (F_1(x), \ldots, F_n(x))$. Изображението $F(x)$ е диферереренцируемо по Фреше́ в $x_0 \in U$ тогава и само тогава, когато $F_i(x)$, $i = 1, \ldots, n$, са диферереренцируеми по Фреше́ в x_0; при това $F'(x_0) = (F_1'(x_0), \ldots, F_n'(x_0))$.*

Доказателство. Нека първо $F_i(x)$, $i = 1, \ldots, n$, са диферереренцируеми по Фреше́ в x_0. Дефинираме изображението $\Lambda : X \to Y$, като полагаме $\Lambda h := (F_1'(x_0)h, \ldots, F_n'(x_0)h)$. Очевидно Λ е непрекъснат (ограничен) линеен оператор. За него имаме с произволно $h \in X$

$$\|F(x_0+h) - F(x_0) - \Lambda h\|_Y$$
$$= \max_{i=1,\dots,n} \|F_i(x_0+h) - F_i(x_0) - F_i'(x_0)h\|_{Y_i} = o(\|h\|_X).$$

Следователно $F(x)$ е дифереренцируемо по Фреше́ в x_0 и $F'(x_0) = \Lambda$.

Обратно, нека $F(x)$ е дифереренцируемо по Фреше́ в x_0. За $i \in \{1,\dots,n\}$ да означим с $\Lambda_i h$ проекцията на $F'(x_0)h$ в Y_i (т.е. $\Lambda_i h$ е i-тата координата на $F'(x_0)h$). Ясно е, че Λ_i е непрекъснат (ограничен) линеен оператор, дефиниран в X със стойности в Y_i.

Имаме

$$\|F_i(x_0+h) - F_i(x_0) - \Lambda_i h\|_{Y_i} \le \|F(x_0+h) - F(x_0) - F'(x_0)h\|_Y = o(\|h\|_X).$$

Следователно $F_i(x)$ е дифереренцируемо по Фреше́ в x_0 и $F_i'(x_0) = \Lambda_i$. $\square$

Нека X е линейно пространство. Интервал (отсечка) в X с краища точките $a, b \in X$, ще наричаме множеството

$$[a,b] := \{x \in X : x = \alpha a + (1-\alpha)b,\ \alpha \in [0,1]\}.$$

Теорема 2.12 (Формула за крайните нараствания). *Нека X и Y са нормирани линейни пространства, $U \subseteq X$ е отворено и $[x_0, x] \subset U$. Нека още изображението $F : U \to Y$ е диференцируемо по Гато́ във всяка точка на $[x_0, x]$. Тогава*

$$\|F(x) - F(x_0)\|_Y \le \sup_{\theta \in [0,1]} \|F_G'(x_0 + \theta(x - x_0))\|_{X \to Y} \|x - x_0\|_X.$$

Доказателство. Полагаме $h := x - x_0$. За $y^* \in Y^*$ разглеждаме реалнозначната функция $\varphi(t) := \langle y^*, F(x_0 + th)\rangle$, $t \in [0,1]$. Тя е диференцируема. Действително за диференчното ѝ частно имаме

$$\frac{\varphi(t+\delta) - \varphi(t)}{\delta} = \left\langle y^*, \frac{F(x_0 + th + \delta h) - F(x_0 + th)}{\delta} \right\rangle.$$

Функционалът y^* е непрекъснат и следователно можем да разместим неговото прилагане с преминаването в граница $\delta \to 0$ в аргумента му, а последната съществува, защото изображението F е диференцируемо по Гато́ във всяка точка на отсечката $[x_0, x]$; при това

$$\lim_{\delta \to 0} \frac{F(x_0 + th + \delta h) - F(x_0 + th)}{\delta} = F_G'(x_0 + th)h.$$

Така получаваме

$$\begin{aligned}\varphi'(t) &= \lim_{\delta\to 0}\frac{\varphi(t+\delta)-\varphi(t)}{\delta}\\ &= \lim_{\delta\to 0}\left\langle y^*, \frac{F(x_0+th+\delta h)-F(x_0+th)}{\delta}\right\rangle\\ &= \left\langle y^*, \lim_{\delta\to 0}\frac{F(x_0+th+\delta h)-F(x_0+th)}{\delta}\right\rangle\\ &= \langle y^*, F'_G(x_0+th)h\rangle.\end{aligned}$$

Към функцията $\varphi(t)$ прилагаме познатата формула за крайните нараствания върху $[0,1]$. Така получаваме, че за някое $\theta\in[0,1]$ (θ изобщо зависи от y^*) имаме

$$\varphi(1)-\varphi(0)=\varphi'(\theta),$$

което означава, че

$$\langle y^*, F(x_0+h)-F(x_0)\rangle = \langle y^*, F'_G(x_0+\theta h)h\rangle.$$

Следователно

$$\begin{aligned}\langle y^*, F(x_0+h)-F(x_0)\rangle &\le \sup_{\theta\in[0,1]}\langle y^*, F'_G(x_0+\theta h)h\rangle\\ &\le \|y^*\|\sup_{\theta\in[0,1]}\|F'_G(x_0+\theta h)\|_{X\to Y}\|h\|_X\end{aligned}\tag{2.2}$$

за всеки ограничен линеен функционал $y^*\in Y^*$.

Накрая да фиксираме такова $y^*\in Y^*$, $y^*\neq 0$, че да имаме

$$\langle y^*, F(x_0+h)-F(x_0)\rangle = \|y^*\|\,\|F(x_0+h)-F(x_0)\|_Y.$$

Такова y^* съществува благодарение на Следствие 1.4. Сега твърдението на теоремата следва от (2.2). □

Като следствие можем да изведем полезно достатъчно условие за дифернцируемост по Фреше́, което е основано на диференцируемост по Гато́.

Следствие 2.13. *Ако $F'_G(x)$ съществува в околност на точката x_0 и е непрекъсната по x в x_0 (в операторната норма), то съществува и $F'(x_0)$, като $F'(x_0)=F'_G(x_0)$.*

Доказателство. Да положим $h=x-x_0$. Прилагаме Теорема 2.12 към изображението $G(x):=F(x)-F'_G(x_0)(x-x_0)$. Така поучаваме неравенството

$$\|F(x_0+h)-F(x_0)-F'_G(x_0)h\|_Y \le \sup_{\theta\in[0,1]}\|F'_G(x_0+\theta h)-F'_G(x_0)\|_{X\to Y}\|h\|_X.$$

След като $F'_G(x)$ съществува в някаква околност на точката x_0 и е непрекъсната по x в x_0 в операторната норма, то за всяко $\varepsilon > 0$ съществува околност на x_0 такава, че за всяко x в нея имаме $\|F'_G(x) - F'_G(x_0)\|_{X\to Y} < \varepsilon$. Следователно съществува $\delta > 0$ такова, че $\sup_{\theta\in[0,1]} \|F'_G(x_0 + \theta h) - F'_G(x_0)\|_{X\to Y} < \varepsilon$, ако $\|h\|_X < \delta$.

Така установихме, че линейният ограничен оператор $F'_G(x_0) : X \to Y$ удовлетворява дефиницията за производна на Фреше. □

2.3 Производни на някои основни типове изображения

(а) Афинни изображения

Изображение $A : X \to Y$ от вида

$$A(x) := \Lambda x + a,$$

където $\Lambda : X \to Y$ е ограничен линеен оператор и $a \in Y$, се нарича афинно. Нека X и Y са нормирани линейни пространства. Непосредствено от дефиницията следва, че A е диференцируемо по Фреше в X, като за всяко $x \in X$

$$A'(x) = \Lambda. \tag{2.3}$$

В частност, ако $x^* \in X^*$ и $a \in \mathbb{R}$, то функционалът

$$f(x) := \langle x^*, x\rangle + a$$

е диференцируем по Фреше в X и $f'(x) = x^*$.

(б) Векторно-значни функции на няколко променливи

Така още ще наричаме изображенията на няколко реални променливи със стойности в крайномерно евклидово пространство. Както вече споменахме, в този случай производната на Фреше съвпада с пълната производна на изображението и познатите теореми от стандартния курс по анализ са приложими. По-точно, нека $U \subseteq \mathbb{R}^n$ е отворено и изображението $F : U \to \mathbb{R}^m$ се задава чрез своите координатни функции посредством

$$F(x) := (f_1(x), \ldots, f_m(x)), \quad x := (x_1, \ldots, x_n).$$

Тогава, ако $f_1, \ldots, f_m$ са диференцируеми частно по всяка една от променливите и частните им производни са непрекъснати в U (накратко, $f_1, \ldots, f_m$ са непрекъснато диференцируеми), то F е диференцируемо по Фреше, като

$$F'(x) = \left(\frac{\partial f_i}{\partial x_j}(x)\right)_{i,j=1}^{m,n}. \tag{2.4}$$

Действието на линейния ограничен оператор $F'(x)$ върху $h \in \mathbb{R}^n$ се интерпретира именно като умножението на матрицата горе вдясно с вектор-стълба h.

В частност, в случая на функция на една реална променлива $f(x)$, производната ѝ на Фреше представлява линеен ограничен функционал от $\mathbb{R}$ в $\mathbb{R}$, чието действие върху $h \in \mathbb{R}$ се задава чрез произведението $f'(x)h$. В този случай производната на Фреше на f в точката x може да се интерпретира като матрица с един единствен елемент ($m = n = 1$).

(в) Оператори върху пространства от функции

(в.1) Разглеждаме ограничения линеен функционал $G : C[t_0, t_1] \to \mathbb{R}$, дефиниран посредством

$$G(x) := \int_{t_0}^{t_1} x(t)\,dt. \tag{2.5}$$

От установеното в (а) следва, че той е диференцируем по Фреше навсякъде, като $G'(x) = G$, т.е. за всяко $x \in C[t_0, t_1]$ имаме

$$G'(x)h = G(h) = \int_{t_0}^{t_1} h(t)\,dt, \quad h \in C[t_0, t_1]. \tag{2.6}$$

(в.2) Нека реално-значната функция $f(t, x)$ е непрекъсната в $[t_0, t_1] \times \mathbb{R}$ и е диференцируема частно по x, като $f'_x(t, x)$ също е непрекъсната в $[t_0, t_1] \times \mathbb{R}$. Разглеждаме изображението $F : C[t_0, t_1] \to C[t_0, t_1]$, дефинирано чрез

$$[F(x)](t) := f(t, x(t)), \quad t \in [t_0, t_1], \quad x \in C[t_0, t_1]. \tag{2.7}$$

Ще докажем, че $F(x)$ е диференцируемо по Фреше навсякъде и ще намерим неговата производна. За тази цел ще приложим Следствие 2.13. Друго директно доказателство е изложено в Допълнението в края на тази глава.

Нека $x, h \in C[t_0, t_1]$ са произволно фиксирани. Първо намираме производната на F по посоката h. Предвид Бележка 2.2 въвеждаме функцията $\varphi(\alpha) := F(x + \alpha h) = f(t, x(t) + \alpha h(t))$. След като $f(t, x)$ е (непрекъснато) диференцируема по x, то $\varphi(\alpha)$ също е диференцируема, като

$$\varphi'(\alpha) = f'_x(t, x(t) + \alpha h(t))h(t).$$

Следователно F е диференцируемо по посоката h, като

$$[F'(x, h)](t) = \varphi'(0) = f'_x(t, x(t))h(t).$$

Очевидно $F'(x, h)$ е линейно изображение по h. Като приложим теоремата на Вайерщрас за ограниченост на непрекъснатите функции върху компакт и използваме непрекъснатостта на f'_x и x, заключаваме, че съществува $c > 0$, такова, че $\sup_{t\in[t_0,t_1]} |f'_x(t, x(t))| \le c$ и следователно

$$\|F'(x, h)\|_C \le c\,\|h\|_C.$$

Тук $\|x\|_C := \max_{x\in[t_0,t_1]} |x(t)|$ е стандартната норма в линейното пространство $C[t_0, t_1]$. С нея то е банахово (пълно).

Така установихме, че операторът $F'(x, h)$ е ограничен относно h. Следователно F е диференцируемо по Гато́ навсякъде, при това

$$[F'_G(x)h](t) = f'_x(t, x(t))h(t).$$

Предвид Следствие 2.13, за да докажем, че F е диференцируемо по Фреше́, остава да се убедим, че изображението $F'_G(x)$ е непрекъснато по x в операторната норма. Но това веднага следва от равномерната непрекъснатост на $f'_x(t, x)$ върху всеки компакт в $[t_0, t_1] \times \mathbb{R}$.[1]

Така установихме, че F е диференцируемо по Фреше́ навсякъде, като

$$[F'(x)h](t) = f'_x(t, x(t))h(t), \quad h \in C[t_0, t_1]. \tag{2.8}$$

Нека по-общо реално-значната функция $f(t, x)$ е непрекъсната в множеството $[t_0, t_1] \times \mathbb{R}^n$, $x := (x_1, \dots, x_n)$, и е диференцируема частно по $x_1, \dots, x_n$, като $f'_{x_i}(t, x)$ също са непрекъснати в $[t_0, t_1]\times\mathbb{R}^n$ за $i = 1, \dots, n$. С $C([t_0, t_1], \mathbb{R}^n)$ ще означаваме банаховото пространство от всички непрекъснати изображения $x : [t_0, t_1] \to \mathbb{R}^n$ с норма $\|x\|_C := \max_{t\in[t_0,t_1]} |x(t)|$, където $|x(t)| := \sqrt{x_1^2(t) + x_2^2(t) + \dots + x_n^2(t)}$, $x := (x_1, \dots, x_n)$, е стандартната евклидова норма в $\mathbb{R}^n$.

Разглеждаме изображението $F : C([t_0, t_1], \mathbb{R}^n) \to C[t_0, t_1]$, дефинирано чрез

$$[F(x)](t) := f(t, x(t)), \quad t \in [t_0, t_1], \quad x \in C([t_0, t_1], \mathbb{R}^n). \tag{2.9}$$

Както по-горе, се доказва, че $F(x)$ е диференцируемо по Фреше́ навсякъде и

$$[F'(x)h](t) = f'_x(t, x(t)) \cdot h(t), \quad h \in C([t_0, t_1], \mathbb{R}^n), \tag{2.10}$$

където $f'_x(t, x)$ е градиентът на $f(t, x)$ по променливата $x := (x_1, \dots, x_n)$.

(в.3) Преминаваме към следващия важен случай. Нека, както погоре, реално-значната функция $f(t, x)$ е непрекъсната в $[t_0, t_1] \times \mathbb{R}$ и е диференцируема частно по x, като $f'_x(t, x)$ също е непрекъсната в $[t_0, t_1]\times \mathbb{R}$. Разглеждаме изображението $H : C[t_0, t_1] \to \mathbb{R}$, дефинирано чрез

$$H(x) := \int_{t_0}^{t_1} f(t, x(t))\, dt.$$

[1] Установяването на това твърдение е добро упражнение.

То може да се представи като съставно изображение $H(x) = G(F(x))$, където F и G са дефинирани съответно в (2.7) и (2.5). От теоремата за диференциране на съставни изображения и формули (2.6) и (2.8) следва, че $H(x)$ е диференцируемо по Фреше́ и

$$H'(x)h = \int_{t_0}^{t_1} f'_x(t, x(t))h(t)\,dt, \quad x, h \in C[t_0, t_1].$$

Нека $f(t,x)$ е непрекъсната в множеството $[t_0,t_1]\times\mathbb{R}^n$, $x := (x_1, \dots, x_n)$, и е диференцируема частно по $x_1, \dots, x_n$, като $f'_{x_i}(t,x)$ също са непрекъснати в $[t_0,t_1] \times \mathbb{R}^n$ за $i = 1, \dots, n$. Съвсем аналогично, посредством (2.6) и (2.10), се установява, че функционалът $H : C([t_0,t_1],\mathbb{R}^n) \to \mathbb{R}$, дефиниран чрез

$$H(x) := \int_{t_0}^{t_1} f(t, x(t))\,dt,$$

е диференцируем по Фреше́ и

$$H'(x)h = \int_{t_0}^{t_1} f'_x(t, x(t)) \cdot h(t)\,dt, \quad x, h \in C([t_0, t_1], \mathbb{R}^n). \tag{2.11}$$

Едно полезно упражнение представлява директният извод на формула (2.11) посредством средствата на диференциалното и интегралното смятане за функции. Това е направено в Допълнението, дадено в края на тази част.

(г) Вектор-функционали върху пространства от функции

Нека $a : \mathbb{R}^n \to \mathbb{R}^m$ е векторно-значна функция на n променливи, задена чрез своите координатни функции

$$a(y) := (a_1(y), \dots, a_m(y)), \quad y := (y_1, \dots, y_n).$$

Нека $\tau \in [t_0, t_1]$ е фиксирано. Разглеждаме изображението

$$H : C([t_0, t_1], \mathbb{R}^n) \to \mathbb{R}^m,$$

дефинирано посредством

$$H(x) := a(x(\tau)), \quad x \in C([t_0, t_1], \mathbb{R}^n).$$

Да предположим, че координатните функции на $a(y)$ са непрекъснато диференцируеми в околност U на точката $y_0 \in \mathbb{R}^n$ и $x_0(\tau) = y_0$. Ще покажем, че H е диференцируемо по Фреше́ в x_0 и ще намерим неговата производна. За да направим това, представяме H като композиция на изображенията $a(y)$ и $F(x) := x(\tau)$, т.е. $H(x) = a(F(x))$. Както знаем, първото е диференцируемо по Фреше́ в y_0, като (вж. (2.4))

$$a'(y_0) = \left(\frac{\partial a_i}{\partial x_j}(y_0)\right)_{i,j=1}^{m,n},$$

а второто е линейно и според (2.3) $F'(x) = F$ за всяко $x \in C([t_0, t_1], \mathbb{R}^n)$, т.е. за всяко $x \in C([t_0, t_1], \mathbb{R}^n)$ имаме $F'(x)h = h(\tau)$, $h \in C([t_0, t_1], \mathbb{R}^n)$. Сега от теоремата за диференциране на съставни изображения (Теорема 2.10) следва, че H е диференцируемо по Фреше́ в x_0, като

$$H'(x_0)h = a'(F(x_0))F'(x_0)h = a'(x_0(\tau))h(\tau), \quad h \in C([t_0, t_1], \mathbb{R}^n). \tag{2.12}$$

Да отбележим изрично, че изразът горе вдясно представлява умножение на матрицата $a'(x_0(\tau)) = a'(y_0)$ от размерност $m \times n$ с вектор-стълба $h(\tau) \in \mathbb{R}^n$. Резултатът от тяхното умножение е вектор-ред в $\mathbb{R}^m$, както трябва да очакваме, след като $H'(x_0)$ е ограничен линеен оператор, действащ от $C([t_0, t_1], \mathbb{R}^n)$ в $\mathbb{R}^m$.

Ще завършим тази част с пример, който показва, че производните на Гато́ и Фреше́ са различни понятия.

Пример 2.1. Нека функцията на две променливи $F(x_1, x_2)$ е дефинирана в равнината посредством

$$F(x_1, x_2) := \begin{cases} 1, & x_1 = x_2^2,\ x_2 > 0, \\ 0 & \text{иначе.} \end{cases}$$

Ще изследваме въпроса за нейната диференцируемост в точката $(0, 0)$. Започваме с производната по посоката $(h_1, h_2) \in \mathbb{R}^2$. Ако $h_1 \leq 0$ или $h_2 \leq 0$, то $F(0 + th_1, 0 + th_2) - F(0, 0) = 0$ за всяко $t > 0$ и следователно $F(x_1, x_2)$ е диференцируема в точката $(0, 0)$ по посоката (h_1, h_2), като $F'((0, 0), (h_1, h_2)) = 0$. Ако пък $h_1, h_2 > 0$, то за достатъчно малки положителни t отново имаме $F(0 + th_1, 0 + th_2) - F(0, 0) = 0$ и следователно $F'((0, 0), (h_1, h_2)) = 0$. Така установихме, че $F(x_1, x_2)$ е диференцируема в точката $(0, 0)$ по всяка посока (h_1, h_2), при това $F'((0, 0), (h_1, h_2)) = 0$ за всеки елемент (h_1, h_2). Оттук следва, че разглежданата функция притежава първа вариация в точката $(0, 0)$ и тя е тъждестено нула. След като тъждествено нулевото изображение представлява непрекъснат линеен оператор (функционал в този случай), то $F(x_1, x_2)$ е диференцируема по Гато́ в точката $(0, 0)$, като $F'_G(0, 0) = (0, 0)$. Тук накрая използвахме, че всеки ограничен линеен функционал, дефиниран в крайномерно пространство, се представя във вида (1.1), като (както лесно се вижда) това представяне е единствено.

Да се върнем на разглежданата функция $F(x_1, x_2)$. Доказахме, че тя е диференцируема по Гато́ в точката $(0, 0)$. Веднага обаче се вижда, че тя е прекъсната в нея. Оттук, както следва от Твърдение 2.7, тя не е диференцируема по Фреше́ в точката $(0, 0)$.

2.4 Теорема на Люстерник

Ще завършим с теоремата на Люстерник. Преди да я формулираме, ще въведем още две основни понятия в диференциалното смятане за оператори.

Дефиниция 2.14. Нека X е банахово пространство, $M \subseteq X$ и $x_0 \in M$. Казваме, че векторът $h \in X$ е *допирателен* към M в точката x_0, ако съществуват $\varepsilon > 0$ и изображение $r : [0, \varepsilon] \to X$ такива, че $x_0 + th + r(t) \in M$ за всяко $t \in [0, \varepsilon]$ и $\|r(t)\|_X = o(t)$.

Множеството от всички допирателни вектори към M в точката x_0 представлява затворен непразен конус. Той се нарича *допирателен конус* на M в точката x_0. Ако той е подпространство, тогава се нарича *допирателно пространство* на M в точката x_0 и се означава чрез $TM(x_0)$. В този случай, т.к. ако $h \in TM(x_0)$, то и $-h \in TM(x_0)$, получаваме, че съществуват $\varepsilon > 0$ и изображение $r : [-\varepsilon, \varepsilon] \to X$ такива, че $x_0 + th + r(t) \in M$ за всяко $t \in [-\varepsilon, \varepsilon]$ и $\|r(t)\|_X = o(t)$.

Теорема 2.15 (Люстерник). *Нека X и Y са банахови пространства, $U \subseteq X$ е отворено и $x_0 \in U$. Нека още $F : U \to Y$ е диференцируемо по Фреше́ в U, като $F'(x)$ е непрекъсната в точката x_0 (в операторната норма) и $\operatorname{Im} F'(x_0) = Y$. Полагаме*

$$M := \{x \in U : F(x) = F(x_0)\}.$$

Тогава

$$TM(x_0) = \operatorname{Ker} F'(x_0);$$

при това съществуват околност $V \subseteq U$ на x_0, константа $c > 0$ и изображение $x : V \to X$ такива, че за всяко $\xi \in V$ имаме

$$F(\xi + x(\xi)) = F(x_0)$$

и

$$\|x(\xi)\|_X \leq c\|F(\xi) - F(x_0)\|_Y.$$

Задачи

1. Докажете, че производната на Фреше́ на даден оператор в дадена вътрешна точка, ако съществува, е единствена.
2. Изследвайте въпроса за диференцируемост по Фреше́ на нормата в хилбертово пространство.
3. Намерете втората производна на Фреше́ на линейно изображение.
4. Докажете, че всяко едно от следните изображения е диференцируемо по Фреше́ във всяка точка от своята дефиниционна област и намерете неговата производна на Фреше́:

(а) $\displaystyle\int_0^1 tx(t)\,dt,\ x \in C[0,1]$;

(б) $\displaystyle\int_0^1 x^2(t)\,dt,\ x \in C[0,1]$;

(в) $\displaystyle\int_0^1 tx^2(t)\,dt,\ x \in C[0,1]$;

(г) $\displaystyle\int_0^1 \dot{x}(t)\,dt,\ x \in C^1[0,1]$;

(д) $\displaystyle\int_0^1 \dot{x}^2(t)\,dt,\ x \in C^1[0,1]$;

(е) $\displaystyle\int_0^1 \big(x^2(t) + \dot{x}^2(t)\big)dt,\ x \in C^1[0,1]$.

5. Докажете Твърдение 2.9.

6. Установете формули за производните по посока и по Гато́ за съставни изображения.

7. В условията на теоремата на Люстерник, докажете, че $TM(x_0) \subseteq \operatorname{Ker} F'(x_0)$. Покажете, че от втората част на твърдението на теоремата (тази след“при това”) следва и обратното включване $\operatorname{Ker}F'(x_0) \subseteq TM(x_0)$.

Допълнение

Тук ще установим диференцируемостта по Фреше и ще намерим производните на някои оператори съвсем непосредствено, без да разбиваме разглежданията на стъпки, както направихме в 2.3.

(д.1) За $x \in C([t_0,t_1],\mathbb{R}^n)$ и $u \in C([t_0,t_1],\mathbb{R}^m)$ разглеждаме функционала

$$\mathcal{F}(x,u) := \int_{t_0}^{t_1} f(t,x(t),u(t))\,dt,$$

където $f \in C^1(\mathbb{R}^{m+n+1})$. За краткост полагаме

$$z := (x,u), \quad g(z) := \mathcal{F}(x,u).$$

Ще докажем, че функционалът g е диференцируем по Фреше в точката $\widetilde{z} = (\widetilde{x},\widetilde{u})$, като изхождаме непосредствено от твърдения на диференциалното и интегралното смятане за функции на няколко променливи. При направените предположения за гладкостта на функцията f, прилагаме формулата на Лайбниц-Нютон и правилото за диференциране на съставни функции на няколко променливи, за да получим за всяко $t \in [t_0, t_1]$ и за произволни функции $h \in C([t_0, t_1], \mathbb{R}^n)$ и $k \in C([t_0, t_1], \mathbb{R}^m)$ такива, че $(h, k) \neq (0, 0)$ и $\|h\|_C$ и $\|k\|_C$ са достатъчно малки,[1] представянето

$$f(t, \widetilde{x}(t) + h(t), \widetilde{u}(t) + k(t)) = f(t, \widetilde{x}(t), \widetilde{u}(t)) +$$
$$\int_0^1 \big[f'_x(t, \widetilde{x}(t) + \tau h(t), \widetilde{u}(t) + \tau k(t)) \cdot h(t) + f'_u(t, \widetilde{x}(t) + \tau h(t), \widetilde{u}(t) + \tau k(t)) \cdot k(t)\big]\, d\tau.$$

Нека изрично отбележим, че горе $f'_x(t, x, u)$ е градиентът на $f(t, x, u)$ по променливата $x = (x_1, x_2, \dots, x_n)$, а $f'_u(t, x, u)$ — градиентът на $f(t, x, u)$ по променливата $u = (u_1, u_2, \dots, u_m)$.

Като положим $\zeta = (h, k)$ и интегрираме последното равенство по t от t_0 до t_1, получаваме

$$\begin{aligned} g(\widetilde{z} + \zeta) &= g(\widetilde{z}) \\ &+ \int_{t_0}^{t_1} \big[f'_x(t, \widetilde{x}(t), \widetilde{u}(t)) \cdot h(t) + f'_u(t, \widetilde{x}(t), \widetilde{u}(t)) \cdot k(t)\big]\, dt + r(\zeta), \end{aligned} \tag{2.13}$$

където сме положили

$$r(\zeta) := \int_{t_0}^{t_1} \int_0^1 \big[f'_x(t, \widetilde{x}(t) + \tau h(t), \widetilde{u}(t) + \tau k(t)) - f'_x(t, \widetilde{x}(t), \widetilde{u}(t))\big] \cdot h(t)\, dt d\tau$$
$$+ \int_{t_0}^{t_1} \int_0^1 \big[f'_u(t, \widetilde{x}(t) + \tau h(t), \widetilde{u}(t) + \tau k(t)) - f'_u(t, \widetilde{x}(t), \widetilde{u}(t))\big] \cdot k(t)\, dt d\tau.$$

Изразът

$$\int_{t_0}^{t_1} \big[f'_x(t, \widetilde{x}(t), \widetilde{u}(t)) \cdot h(t) + f'_u(t, \widetilde{x}(t), \widetilde{u}(t)) \cdot k(t)\big]\, dt$$

дефинира един ограничен линеен функционал на ζ в банаховото пространство $C([t_0, t_1], \mathbb{R}^n) \times C([t_0, t_1], \mathbb{R}^m)$ с норма $\|\zeta\|_C := \max\{\|h\|_C, \|k\|_C\}$. За да покажем, че $g(z)$ е диференцируемо по Фреше в $\widetilde{z}$ и този израз представлява $g'(\widetilde{z})\zeta$, остава да покажем, че $r(\zeta) = o(\|\zeta\|_C)$.

Да означим двойните интеграли в $r(\zeta)$ съответно с $r_x(\zeta)$ и $r_u(\zeta)$. За $r_x(\zeta)$ имаме

$$|r_x(\zeta)| \leq \int_{t_0}^{t_1} \int_0^1 \big|[f'_x(t, \widetilde{x}(t) + \tau h(t), \widetilde{u}(t) + \tau k(t)) - f'_x(t, \widetilde{x}(t), \widetilde{u}(t))] \cdot h(t)\big|\, dt d\tau.$$

[1] Искаме $\|h\|_C$ и $\|k\|_C$ да са толкова малки, че да имаме $(t, \widetilde{x}(t) + h(t), \widetilde{u}(t) + k(t)) \in U$ за всяко $t \in [t_0, t_1]$; това е възможно, защото $\widetilde{x}$ и $\widetilde{u}$ са непрекъснати, $(t, \widetilde{x}(t), \widetilde{u}(t)) \in U$ за всяко t и U е отворено в $\mathbb{R}^{n+m+1}$.

Неравенството на Коши влече

$$\begin{aligned}
&\left|[f'_x(t,\widetilde{x}(t)+\tau h(t),\widetilde{u}(t)+\tau k(t))-f'_x(t,\widetilde{x}(t),\widetilde{u}(t))]\cdot h(t)\right|\\
&\qquad\le |f'_x(t,\widetilde{x}(t)+\tau h(t),\widetilde{u}(t)+\tau k(t))-f'_x(t,\widetilde{x}(t),\widetilde{u}(t))|\,|h(t)|\\
&\qquad\le \sup_{\substack{t\in[t_0,t_1]\\ |\xi|,|\eta|\le\|\zeta\|_C}} |f'_x(t,\widetilde{x}(t)+\xi,\widetilde{u}(t)+\eta)-f'_x(t,\widetilde{x}(t),\widetilde{u}(t))|\,\|\zeta\|_C.
\end{aligned}$$

Да забележим, че точките $(t,\widetilde{x}(t)+\xi,\widetilde{u}(t)+\eta)$ описват едно компактно множество в $\mathbb{R}^{n+m+1}$, когато $t\in[t_0,t_1]$ и $|\xi|,|\eta|\le\|\zeta\|_C$. Тогава, щом функцията $f'_x(t,x,u)$ е непрекъсната, то тя е и равномерно непрекъсната върху всеки компакт. Следователно

$$\sup_{\substack{t\in[t_0,t_1]\\ |\xi|,|\eta|\le\|\zeta\|_C}} |f'_x(t,\widetilde{x}(t)+\xi,\widetilde{u}(t)+\eta)-f'_x(t,\widetilde{x}(t),\widetilde{u}(t))|\to 0 \quad \text{при} \quad \|\zeta\|_C\to 0.$$

Така установихме, че $r_x(\zeta)=o(\|\zeta\|_C)$. Съвсем аналогично се вижда, че $r_u(\zeta)=o(\|\zeta\|_C)$. Така имаме $r(\zeta)=o(\|\zeta\|_C)$ и сега от (2.13) следва, че $g(z)$ е диференцируемо по Фреше в $\widetilde{z}$ и

$$\begin{aligned}
g'(\widetilde{z})\zeta=\int_{t_0}^{t_1}[f'_x(t,\widetilde{x}(t),\widetilde{u}(t))\cdot h(t)+f'_u(t,\widetilde{x}(t),\widetilde{u}(t))\cdot k(t)]\,dt,&\\
\zeta=(h,k)\in C([t_0,t_1],\mathbb{R}^n)\times C([t_0,t_1],\mathbb{R}^m).&
\end{aligned}$$

(д.2) С $C^k([t_0,t_1],\mathbb{R}^n)$ ще означаваме банаховото пространство от всички k пъти непрекъснато диференцируеми изображения $x:[t_0,t_1]\to\mathbb{R}^n$ с норма $\|x\|_{C^k}:=\max_{j=0,\dots,k}\|x^{(j)}\|_C$. Нека $\varphi\in C^1(\mathbb{R}^{m+n+1})$. Разглеждаме оператора

$$F: C^1([t_0,t_1],\mathbb{R}^n)\times C([t_0,t_1],\mathbb{R}^m)\to C([t_0,t_1],\mathbb{R}^n),$$

дефиниран чрез

$$F(z):=\dot{x}-\varphi(t,x,u),$$

където с $\dot{x}$ сме означили производната на вектор-функцията $x(t)$, $z:=(x,u)$, $\|z\|_Z:=\max\{\|x\|_{C^1},\|u\|_C\}$.

Нека $h\in C^1([t_0,t_1],\mathbb{R}^n)$ и $k\in C([t_0,t_1],\mathbb{R}^m)$ са такива, че $\|h\|_C$ и $\|k\|_C$ са достатъчно малки. За $\zeta=(h,k)$ имаме

$$\begin{aligned}
F(z+\zeta)&=\dot{x}+\dot{h}-\varphi(t,x+h,u+k)\\
&=[\dot{x}-\varphi(t,x,u)]+\dot{h}-[\varphi(t,x+h,u+k)-\varphi(t,x,u)]\\
&=F(z)+\dot{h}-\int_0^1\left[\varphi'_x(t,x+\tau h,u+\tau k)\,h+\varphi'_u(t,x+\tau h,u+\tau k)\,k\right]d\tau\\
&=F(z)+[\dot{h}-\varphi'_x(t,x,u)\,h-\varphi'_u(t,x,u)\,k]-r_1(\zeta),
\end{aligned}$$

където сме положили

$$r_1(\zeta) := \int_0^1 [\varphi'_x(t, x+\tau h, u+\tau k) - \varphi'_x(t,x,u)]\, h\, d\tau \\ + \int_0^1 [\varphi'_u(t, x+\tau h, u+\tau k) - \varphi'_u(t,x,u)]\, k\, d\tau.$$

Да припомним, че $\varphi'_x(t,x,u)$ е пълната производна (производната на Фреше) на изображението φ по променливата $x = (x_1, x_2, \ldots, x_n)$ в точката (t,x,u). Тя се представя посредством матрица $n \times n$, чиито компоненти са частните производни на координатните функции на φ по $x_1, x_2, \ldots, x_n$ в тази точка. Аналогично $\varphi'_u(t,x,u)$ е пълната производна (производната на Фреше) на изображението φ по променливата $u = (u_1, u_2, \ldots, u_m)$ в точката (t,x,u). Тя се представя посредством матрица $n \times m$, чиито компоненти са частните производни на координатните функции на φ по $u_1, u_2, \ldots, u_m$ в тази точка. Подинтегралните функции горе са векторно-значни. Интегралът от такава функция се разбира като вектора от интегралите на координатните функции.

Както по-горе в (д.1), се убеждаваме, че $r_1(\zeta) = o(\|\zeta\|_Z)$, т.е. $\|r_1(\zeta)\|_C = o(\|\zeta\|_Z)$. Следователно $F(z)$ е диференцируемо по Фреше, при това

$$F'(z)\zeta = \dot{h} - \varphi'_x(t,x,u)\, h - \varphi'_u(t,x,u)\, k, \\ \zeta = (h,k) \in Z = C^1([t_0,t_1], \mathbb{R}^n) \times C([t_0,t_1], \mathbb{R}^m).$$

3 Задача със закрепени краища

Тук ще разгледаме една от основните екстремални задачи на т.нар. вариационно смятане. Нека $U \subseteq \mathbb{R}^3$ е отворено и $L \in C^1(U)$. Нека още $t_0, t_1 \in \mathbb{R}$, $t_0 < t_1$. Търсим минимум на функционала $\mathcal{F} : C^1[t_0, t_1] \to \mathbb{R}$, дефиниран чрез

$$\mathcal{F}(x) := \int_{t_0}^{t_1} L(t, x(t), \dot{x}(t))\, dt$$

при условията

$$x(t_0) = x_0, \quad x(t_1) = x_1.$$

С $\dot{x}(t)$ означаваме производната на $x(t)$, а x_0 и x_1 са дадени реални числа.

Задачи от този тип ще формулираме по-накратко така

$$(3.1) \qquad \left| \begin{aligned} &\mathcal{F}(x) := \int_{t_0}^{t_1} L(t, x(t), \dot{x}(t))\, dt \longrightarrow \inf \\ &x(t_0) = x_0, \quad x(t_1) = x_1. \end{aligned} \right.$$

Те носят наименованието *задачи със закрепени краища*. Ще установим необходимо условие за локален минимум в топологията (нормата) на пространството $C^1[t_0, t_1]$. Да припомним, че то, снабдено с нормата

$$\|x\|_{C^1} := \max\{\|x\|_C, \|\dot{x}\|_C\},$$

където $\|x\|_C := \max_{t \in [t_0, t_1]} |x(t)|$, е банахово пространство.

Следва точната дефиниция.

Дефиниция 3.1. Казваме, че функцията $\widetilde{x} \in C^1[t_0, t_1]$, $\widetilde{x}(t_0) = x_0$, $\widetilde{x}(t_1) = x_1$ е точка на *слаб локален минимум* на задачата (3.1), ако съществува $\delta > 0$ такова, че $\mathcal{F}(\widetilde{x}) \le \mathcal{F}(x)$ за всяко $x \in C^1[t_0, t_1]$ такова, че $\|x - \widetilde{x}\|_{C^1} < \delta$ и $x(t_0) = x_0$, $x(t_1) = x_1$.

За сравнение дефиницията за силен локален минимум гласи

Дефиниция 3.2. Казваме, че функцията $\widetilde{x} \in C^1[t_0, t_1]$, $\widetilde{x}(t_0) = x_0$, $\widetilde{x}(t_1) = x_1$ е точка на *силен локален минимум* на задачата (3.1), ако съществува $\delta > 0$ такова, че $\mathcal{F}(\widetilde{x}) \le \mathcal{F}(x)$ за всяко $x \in C^1[t_0, t_1]$ такова, че $\|x - \widetilde{x}\|_C < \delta$ и $x(t_0) = x_0$, $x(t_1) = x_1$.

Когато е налице ситуацията от горните дефиниции, още ще казваме, че функционалът $\mathcal{F}$ има слаб, съответно силен, локален минимум в $\widetilde{x} \in C^1[t_0, t_1]$ при условията $x(t_0) = x_0$, $x(t_1) = x_1$.

Разликата между двете дефиниции се съдържа в типа околност. В първата тя е в топологията на C^1, а във втората — в топология на C, в която околностите са съществено по-широки. По-точно, т.к. от $\|x - \widetilde{x}\|_{C^1} < \delta$ следва $\|x - \widetilde{x}\|_C < \delta$, то околността на $\widetilde{x}$, определена с условието $\|x - \widetilde{x}\|_{C^1} < \delta$, се съдържа в околността, определената чрез условието

$\|x-\widetilde{x}\|_C < \delta$. При това последната съдържа значително повече функции от първата. Също така, това показва, че всяка точка на силен минимум е и точка на слаб.

Основният резултат, който ще докажем тук е следното необходимо условие за слаб локален минимум.

Теорема 3.3 (Необходимо условие на Ойлер)**.** *Нека $U \subseteq \mathbb{R}^3$ е отворено и $L \in C^1(U)$. Ако $\widetilde{x}(t)$ е слаб локален минимум на задачата* (3.1)*, то*

$$-\frac{d}{dt}L'_{\dot{x}}(t,\widetilde{x}(t),\dot{\widetilde{x}}(t)) + L'_x(t,\widetilde{x}(t),\dot{\widetilde{x}}(t)) = 0, \quad t \in [t_0,t_1]. \tag{3.2}$$

Тук L'_x и $L'_{\dot{x}}$ означават частните производни на L съответно по втората и третата променлива. В твърдението на теоремата неявно се вклюва и това, че функцията $L'_{\dot{x}}(t,\widetilde{x}(t),\dot{\widetilde{x}}(t))$ е диференцируема по t.

Диференциалното уравнение

$$-\frac{d}{dt}L'_{\dot{x}}(t,x(t),\dot{x}(t)) + L'_x(t,x(t),\dot{x}(t)) = 0, \quad t \in [t_0,t_1]. \tag{3.3}$$

се нарича *уравнение на Ойлер-Лагранж*. Неговите решения се наричат *екстремали* на задачата (3.1). Теорема 3.3 показва, че ако задачата (3.1) има решение, то то удовлетворява диференциалното уравнение (3.2) с гранични условия $x(t_0) = x_0$, $x(t_1) = x_1$.

В доказателството на теоремата ще използваме следното помощно твърдение.

Лема 3.4 (Лема на дю Боа-Реймон)**.** *Нека $b \in C[t_0,t_1]$. Ако*

$$\int_{t_0}^{t_1} b(t)v(t)\,dt = 0$$

за всяка функция $v \in C[t_0,t_1]$ такава, че

$$\int_{t_0}^{t_1} v(t)\,dt = 0,$$

то $b(t) \equiv \mathrm{const}$.

Доказателство. Допускаме противното, а именно, че съществуват две точки $\tau_1, \tau_2 \in (t_0,t_1)$, $\tau_1 < \tau_2$, такива, че $b(\tau_1) \neq b(\tau_2)$. Ще предположим, че $b(\tau_1) > b(\tau_2)$. Другият случай се свежда към този, като разгледаме функцията $-b$ вместо b.

Фиксираме $\varepsilon > 0$ толкова малко, че

(а) интервалите $\Delta_1 := [\tau_1-\varepsilon,\tau_1+\varepsilon]$ и $\Delta_2 := [\tau_2-\varepsilon,\tau_2+\varepsilon]$ не се пресичат взаимно и се съдържат в $[t_0,t_1]$;

(б) $b_1 := \min_{t\in\Delta_1} b(t) > b_2 := \max_{t\in\Delta_2} b(t)$.

Разглеждаме функцията $v \in C[t_0, t_1]$, дефинирана по следния начин:

$$v(t) := \begin{cases} 1, & t = \tau_1, \\ -1, & t = \tau_2, \\ 0, & t \notin \Delta_1 \cup \Delta_2, \\ \text{линейна върху} & [\tau_1 - \varepsilon, \tau_1],\ [\tau_1, \tau_1 + \varepsilon],\ [\tau_2 - \varepsilon, \tau_2],\ [\tau_2, \tau_2 + \varepsilon]. \end{cases}$$

Ясно е, че

$$\int_{t_0}^{t_1} v(t)\, dt = 0.$$

С така дефинираната функция $v(t)$ имаме

$$\begin{aligned} \int_{t_0}^{t_1} b(t)v(t)\, dt &= \int_{\Delta_1} b(t)v(t)\, dt + \int_{\Delta_2} b(t)v(t)\, dt \\ &\geq (b_1 - b_2) \int_{\Delta_1} v(t)\, dt > 0, \end{aligned}$$

което е противоречие. □

Преминаваме към доказателството на теоремата.

Доказателство на Теорема 3.3. Нека $h \in C^1[t_0, t_1]$. За $\alpha \in \mathbb{R}$ толкова малко по абсолютна стойност, че да имаме $(t, \widetilde{x}(t) + \alpha h(t), \dot{\widetilde{x}}(t) + \alpha \dot{h}(t)) \in U$, $t \in [t_0, t_1]$, дефинираме функцията

$$\varphi(\alpha) := \mathcal{F}(\widetilde{x} + \alpha h) = \int_{t_0}^{t_1} L(t, \widetilde{x}(t) + \alpha h(t), \dot{\widetilde{x}}(t) + \alpha \dot{h}(t))\, dt.$$

От теоремата за диференциране под знака на интеграла следва, че $\varphi(\alpha)$ е диференцируема и

$$\begin{aligned} \varphi'(\alpha) = \int_{t_0}^{t_1} \Big[& L'_x(t, \widetilde{x}(t) + \alpha h(t), \dot{\widetilde{x}}(t) + \alpha \dot{h}(t))\, h(t) \\ & + L'_{\dot{x}}(t, \widetilde{x}(t) + \alpha h(t), \dot{\widetilde{x}}(t) + \alpha \dot{h}(t))\, \dot{h}(t) \Big] dt. \end{aligned}$$

Следователно

$$\begin{aligned} \mathcal{F}'(\widetilde{x}, h) &= \varphi'(0) \\ &= \int_{t_0}^{t_1} \Big[L'_x(t, \widetilde{x}(t), \dot{\widetilde{x}}(t))\, h(t) + L'_{\dot{x}}(t, \widetilde{x}(t), \dot{\widetilde{x}}(t))\, \dot{h}(t) \Big]\, dt. \end{aligned} \tag{3.4}$$

Да предположим още, че $h(t_0) = h(t_1) = 0$. Тогава, т.к. функционалът $\mathcal{F}$ има слаб локален минимум в $\widetilde{x}$, то $\varphi(\alpha)$ има локален минимум в $\alpha = 0$. От теоремата на Ферма сега следва, че $\varphi'(0) = 0$. Така установихме, че

$$\mathcal{F}'(\widetilde{x}, h) = 0 \quad \forall h \in C_0^1[t_0, t_1], \tag{3.5}$$

където
$$C_0^1[t_0,t_1] := \{x \in C^1[t_0,t_1] : x(t_0) = x(t_1) = 0\}.$$

По-нататък, да положим

$$p(t) := L'_{\dot{x}}(t, \widetilde{x}(t), \dot{\widetilde{x}}(t)), \quad q(t) := L'_x(t, \widetilde{x}(t), \dot{\widetilde{x}}(t)).$$

Формула (3.4) придобива вида

$$\int_{t_0}^{t_1} [q(t)\, h(t) + p(t)\, \dot{h}(t)]\, dt = 0 \quad \forall h \in C_0^1[t_0,t_1]. \tag{3.6}$$

Интегрираме по части интеграла от $q(t)\, h(t)$. Получаваме

$$\begin{aligned} \int_{t_0}^{t_1} q(t)\, h(t)\, dt &= \int_{t_0}^{t_1} h(t)\, d\left(\int_{t_1}^{t} q(\tau)\, d\tau\right) \\ &= \int_{t_0}^{t_1} Q(t)\, \dot{h}(t)\, dt, \end{aligned} \tag{3.7}$$

където сме положили
$$Q(t) := \int_t^{t_1} q(\tau)\, d\tau.$$

Като комбинираме (3.6) и (3.7), получаваме

$$\int_{t_0}^{t_1} [Q(t) + p(t)]\, \dot{h}(t)\, dt = 0 \quad \forall h \in C_0^1[t_0,t_1]. \tag{3.8}$$

Сега от лемата на дю Боа-Реймон с $b(t) = Q(t) + p(t)$ следва, че

$$Q(t) + p(t) \equiv \text{const}, \quad t \in [t_0,t_1]. \tag{3.9}$$

Тази лема е приложима, защото всяка непрекъсната функция $v(t)$ върху $[t_0,t_1]$ такава, че
$$\int_{t_0}^{t_1} v(t)\, dt = 0,$$
се представя във вида $v(t) = \dot{h}(t)$ с $h \in C_0^1[t_0,t_1]$. Действително, това е така с
$$h(t) := \int_{t_0}^{t} v(\tau)\, d\tau.$$

За да приключим доказателството, остава само да отбележим, че от (3.9) следва $p \in C^1[t_0,t_1]$ и

$$-\dot{p}(t) + q(t) = 0, \quad t \in [t_0,t_1],$$

което е точно (3.2). □

Ще завършим с няколко примера.

Пример 3.1. Разглеждаме задачата със закрепени краища

$$\mathcal{F}(x) := \int_0^1 \dot{x}(t)^2 dt \longrightarrow \inf, \quad x(0) = x(1) = 0, \quad x \in C^1[0,1].$$

Тук $L(t, x, \dot{x}) = \dot{x}^2$. Тогава $L'_x = 0$ и $L'_{\dot{x}} = 2\dot{x}$. Уравнението на Ойлер-Лагранж (3.3) има вида $(d/dt)(2\dot{x}) = 0$, т.е. $\ddot{x} = 0$. Неговите решения (екстремалите на задачата) представляват линейните функции. От тях единствено тъждествено нулевата функция удовлетворява граничните условия $x(0) = x(1) = 0$. Така установихме, че ако задачата има решение, то то е $\widetilde{x}(t) \equiv 0$. Очевидно тази функция е решение, при това в силен смисъл и глобално, защото $\mathcal{F}(x) \geq 0$ за всяко $x \in C^1[0,1]$.

Пример 3.2. Да разглеждаме сега задачата

$$\mathcal{F}(x) := \int_0^1 \dot{x}(t)^3 dt \longrightarrow \inf, \quad x(0) = 0,\ x(1) = 1, \quad x \in C^1[0,1].$$

Уравнението на Ойлер-Лагранж има вида $(d/dt)(3\dot{x}^2) = 0$. Това влече $\dot{x}^2 = \text{const}$, което предвид непрекъснатостта на $\dot{x}$, е еквивалентно на $\dot{x} = \text{const}$. Така отново получаваме, че екстремалите са линейните функции. Сред тях, тази, която удовлетворява граничните условия, е $\widetilde{x}(t) = t$. Това означава, че ако разглежданата задача има решение в слаб смисъл (а следователно и ако има решение в силен смисъл), то е $\widetilde{x}(t) = t$. За да проверим дали действително то е решение, разглеждаме действието на функционала $\mathcal{F}$ върху функции от вида $\widetilde{x} + h$, където $h \in C_0^1[0,1]$. Имаме

$$\begin{aligned}\mathcal{F}(\widetilde{x} + h) &= \int_0^1 \left(\frac{d}{dt}(t + h(t))\right)^3 dt = \\ &= \mathcal{F}(\widetilde{x}) + 3\int_0^1 \dot{h}(t)\, dt + \int_0^1 \left(3\dot{h}(t)^2 + \dot{h}(t)^3\right) dt.\end{aligned}$$

Отчитайки, че $h(0) = h(1) = 0$, получаваме

$$\int_0^1 \dot{h}(t)\, dt = 0$$

и следователно

$$\mathcal{F}(\widetilde{x} + h) = \mathcal{F}(\widetilde{x}) + \int_0^1 \dot{h}(t)^2 \left(3 + \dot{h}(t)\right) dt. \tag{3.10}$$

Така, ако $h \in C_0^1[0,1]$ е такава, че $\|h\|_{C^1} < 3$, то интегралът горе вдясно е неотрицателен и следователно

$$\mathcal{F}(\widetilde{x} + h) \geq \mathcal{F}(\widetilde{x})$$

и функцията $\widetilde{x}(t) = t$ е точка на слаб локален минимум на функционала $\mathcal{F}$.

Да проверим дали тази функция представлява решение и в силен смисъл. Формулата (3.10) ни подсказва, че това едва ли е така. Тя ни помага и да конструираме контрапример. Ще дефинираме редица $\{h_n\}$ от функции в $C_0^1[0,1]$ такива, че $\|h_n\|_C \to 0$, но същевременно $\int_0^1 \dot{h}_n(t)^3 dt \to -\infty$ и то толкова бързо, че да не може да бъде компенсирано от величината $\int_0^1 \dot{h}_n(t)^2 dt$. Една редица с тези свойства е, например, следната

$$h_n(t) := \int_0^t g_n(\tau)\, d\tau, \quad g_n(\tau) := \begin{cases} -\sqrt{n}, & \tau \in [0, 1/n], \\ j_n(\tau), & \tau \in [1/n, 2/n], \\ \frac{\sqrt{n}}{n-2}, & \tau \in [2/n, 1], \end{cases}$$

където $n \geq 3$, а $j_n(\tau)$ е непрекъсната строго монотонно растяща функция такава, че

$$\int_{1/n}^{2/n} j_n(\tau)\, d\tau = 0, \quad j_n(1/n) = -\sqrt{n}, \quad j_n(2/n) = \frac{\sqrt{n}}{n-2}.$$

Съвсем лесно се вижда, че такава функция j_n съществува за всяко естествено число $n \geq 3$.

Очевидно за така конструираните функции $h_n(t)$ имаме, че са непрекъснато диференцируеми в $[0,1]$ и се анулират в краищата на интервала. По-нататък за техните производни имаме

$$\begin{aligned} \int_0^1 \dot{h}_n(\tau)^3 d\tau &= -\int_0^{1/n} n^{3/2} d\tau + \int_{1/n}^{2/n} j_n(\tau)^3 d\tau + \int_{2/n}^1 \left(\frac{\sqrt{n}}{n-2}\right)^3 d\tau \\ &\leq -\sqrt{n} + \frac{1}{n}\left(\frac{\sqrt{n}}{n-2}\right)^3 + \left(1 - \frac{2}{n}\right)\left(\frac{\sqrt{n}}{n-2}\right)^3 \\ &\leq -\sqrt{n} + \left(\frac{\sqrt{n}}{n-2}\right)^3 \end{aligned}$$

и аналогично

$$\int_0^1 \dot{h}_n(\tau)^2 d\tau \leq 1 + \left(\frac{\sqrt{n}}{n-2}\right)^2.$$

Следователно

$$\mathcal{F}(\widetilde{x} + h_n) \leq -\sqrt{n} + c, \quad n \geq 3, \tag{3.11}$$

с някаква положителна константа c.

Остава да се убедим, че $\|h_n\|_C \to 0$. За всяко $n \geq 3$ функцията h_n е изпъкнала[1] и неположителна. Производната ѝ $\dot{h}_n = g_n$ се анулира в

[1] Това е така, защото производната ѝ е монотонно растяща.

единствена точка. Тя се намира в $(1/n, 2/n)$. Да я означим с ξ_n. Имаме

$$\begin{aligned}\|h_n\|_C &= -h_n(\xi_n) = -\int_0^{\xi_n} g_n(\tau)\,d\tau \\ &= \int_0^{1/n} \sqrt{n}\,d\tau - \int_{1/n}^{\xi_n} j_n(\tau)\,d\tau \\ &\le \frac{1}{\sqrt{n}} + \left(\xi_n - \frac{1}{n}\right)\sqrt{n} \\ &\le \frac{2}{\sqrt{n}} \to 0, \quad n \to \infty.\end{aligned}$$

Така установихме, че $\widetilde{x}(t) = t$ не е точка на силен локален минимум на функционала $\mathcal{F}$. Следователно задата няма силно решение, при това, както се вижда от (3.11), $\inf\{\mathcal{F}(x) : x \in C^1[0,1],\ x(0) = 0,\ x(1) = 1\} = -\infty$.

Пример 3.3. Ще приведем и пример на функционал от разглеждания тип, който е ограничен отдолу, но не съществува функция, отговаряща на условията, която реализира неговата точна долна граница. Този пример е даден от Вайерщрас. Разглеждаме задачата със закрепени краища:

$$\mathcal{F}(x) := \int_0^1 t^2\dot{x}(t)^2 dt \longrightarrow \inf, \quad x(0) = 0,\ x(1) = 1, \quad x \in C^1[0,1].$$

Уравнението на Ойлер-Лагранж има вида $(d/dt)(2t^2\dot{x}) = 0$. Неговото общо решение е $x(t) = c_1t^{-1}+c_2$, $c_1, c_2 \in \mathbb{R}$. Никоя функция от този вид не удовлетворява граничните условия (всъщност, за да имаме $x \in C^1[0,1]$, трябва $c_1 = 0$). От Теорема 3.3 следва, че задачата няма решение в класа $C^1[0,1]$. Може да забележим, че за всяко $x \in C^1[0,1]$ имаме $\mathcal{F}(x) > 0$, докато $\inf\{\mathcal{F}(x) : x \in C^1[0,1],\ x(0) = 0,\ x(1) = 1\}$=0. Вайерщрас предлага минимизиращата редица

$$x_n(t) := \frac{\operatorname{arctg} nt}{\operatorname{arctg} n}.$$

Тези функции удовлетворяват граничните условия и са непрекъснато диференцируеми. Установява се, че

$$\mathcal{F}(x_n) = \frac{1}{\operatorname{arctg}^2 n}\int_0^1 \frac{n^2t^2}{(1+n^2t^2)^2}\,dt \to 0, \quad n \to \infty.$$

Това може да се покаже с помощта на теоремата на Лебег за граничен преход под интеграла. Действително, подинтегралните функции са сумириеми, неотрицателни са, мажорират се от 1 и във всяка точка $t \in [0,1]$ клонят към 0 при $n \to \infty$.[1]

[1] Забележете, че сходимостта не е равномерна и теоремата за граничен преход под интеграла, позната ни от курса по анализ относно римановия интеграл, не е приложима.

Между другото за самата редица $\{x_n\}$ имаме

$$\lim_{n\to\infty} x_n(t) = \begin{cases} 0, & t = 0, \\ 1, & t \in (0,1]. \end{cases}$$

Задачи

1. Нека $U \subseteq \mathbb{R}^3$ е отворено и $L \in C^1(U)$. Докажете, че ако $\widetilde{x} \in C^1[t_0, t_1]$ е слаб локален минимум на задачата със свободен край

$$\int_{t_0}^{t_1} L(t, x(t), \dot{x}(t))\, dt \longrightarrow \inf, \quad x(t_0) = x_0, \quad x \in C^1[t_0, t_1],$$

то

$$\begin{aligned} &-\frac{d}{dt} L'_{\dot{x}}(t, \widetilde{x}(t), \dot{\widetilde{x}}(t)) + L'_x(t, \widetilde{x}(t), \dot{\widetilde{x}}(t)) = 0, \quad t \in [t_0, t_1], \\ &L'_{\dot{x}}(t_1, \widetilde{x}(t_1), \dot{\widetilde{x}}(t_1)) = 0. \end{aligned}$$

Упътване: Твърдението може да се докаже съвсем аналогично на Теорема 3.3.

2. Нека $U \subseteq \mathbb{R}^3$ и $V \subseteq \mathbb{R}$ са отворени. Нека още $L \in C^1(U)$ и $\psi_0, \psi_1 \in C^1(V)$. Докажете, че ако $\widetilde{x} \in C^1[t_0, t_1]$ е слаб локален минимум на задачата на Болца

$$\psi_0(x(t_0)) + \psi_1(x(t_1)) + \int_{t_0}^{t_1} L(t, x(t), \dot{x}(t))\, dt \longrightarrow \inf, \quad x \in C^1[t_0, t_1],$$

то

$$\begin{aligned} &-\frac{d}{dt} L'_{\dot{x}}(t, \widetilde{x}(t), \dot{\widetilde{x}}(t)) + L'_x(t, \widetilde{x}(t), \dot{\widetilde{x}}(t)) = 0, \quad t \in [t_0, t_1], \\ &L'_{\dot{x}}(t_0, \widetilde{x}(t_0), \dot{\widetilde{x}}(t_0)) = \psi'_0(\widetilde{x}(t_0)), \\ &L'_{\dot{x}}(t_1, \widetilde{x}(t_1), \dot{\widetilde{x}}(t_1)) = -\psi'_1(\widetilde{x}(t_1)). \end{aligned}$$

Бележка: Необходимите условия в горните две задачи също носят името на Ойлер.

3. Решете екстремалните задачи (навсякъде $x \in C^1$ в съответния интервал):

   (а) $\displaystyle\int_0^1 \left(\dot{x}^2 + x\right) dt \longrightarrow \inf,\ x(0) = 0,\ x(1) = 1;$

   (б) $\displaystyle\int_1^2 \left(t\dot{x}^2 - x\right) dt \longrightarrow \inf,\ x(1) = 0,\ x(2) = 1;$

   (в) $\displaystyle\int_0^1 \left(\dot{x}^2 + x^2\right) dt - x^2(1) \longrightarrow \inf,\ x(0) = 1.$

4 Гладки задачи. Множители на Лагранж

Ще разгледаме един широк клас екстремални задачи — т.нар. гладки задачи. Наименованието им идва от предположенията за диференцируемост, които се правят върху участващите изображения. Ще докажем обобщенията в случая на изображения, действащи в банахови пространства, на две класически необходими условия за екстремум: условието на Ферма за безусловен локален екстремум и условието на Лагранж за условен локален екстремум.

Първо ще се спрем на необходимо условие за екстремум в случая на липса на допълнителни ограничения.

Дефиниция 4.1. Нека X е линейно топологично пространство и $U \subseteq X$. Казваме, че изображението $f : U \to \mathbb{R}$ има локален минимум в точката $\widetilde{x} \in \operatorname{int} U$ (или още, че $\widetilde{x}$ е точка на локален минимум на изображението $f : U \to \mathbb{R}$), ако съществува околност $V \subseteq U$ на $\widetilde{x}$ такава, че $f(x) \geq f(\widetilde{x})$ за всяко $x \in V$.

Когато X е нормирано, горната дефиниция се изказва още така:

Дефиниция 4.2. Нека X е нормирано линейно пространство и $U \subseteq X$. Казваме, че изображението $f : U \to \mathbb{R}$ има *локален минимум* в точката $\widetilde{x} \in \operatorname{int} U$, ако съществува $\delta > 0$ такова, че $f(x) \geq f(\widetilde{x})$ за всяко $x \in U$ такова, че $\|x - \widetilde{x}\|_X < \delta$.

Аналогично се дефинира понятието за *локален максимум.* Както в диференциалното смятане за функции, общо локалните минимуми и максимуми се наричат *локални екстремуми.*

Твърдение 4.3. *Нека X е линейно топологично пространство и $U \subseteq X$. Ако изображението $f : U \to \mathbb{R}$ има локален минимум в точката $\widetilde{x}$ и производната по посока $f'(\widetilde{x}, h)$ съществува, то $f'(\widetilde{x}, h) \geq 0$.*

Доказателство. След като $f : U \to \mathbb{R}$ има локален минимум в точката $\widetilde{x}$, то за достатъчно малки $t > 0$ имаме

$$\frac{f(\widetilde{x} + th) - f(\widetilde{x})}{t} \geq 0.$$

След граничен преход $t \to 0+$, получаваме

$$f'(\widetilde{x}, h) = \lim_{t \to 0+} \frac{f(\widetilde{x} + th) - f(\widetilde{x})}{t} \geq 0.$$

$\square$

От това твърдение непосредствено се извежда следното обобщение на необходимото условие на Ферма за локален екстремум.

Теорема 4.4 (Обобщена теорема на Ферма)**.** *Нека X е нормирано линейно пространство и $U \subseteq X$. Ако изображението $f : U \to \mathbb{R}$ има локален екстремум в точката $\widetilde{x}$ и производната му на Гато $f'_G(\widetilde{x})$ съществува, то $f'_G(\widetilde{x}) = 0$. В частност, твърдението е вярно за производна на Фреше.*

Преминаваме към задачата за условен локален екстремум.

Дефиниция 4.5. Нека X и Y са линейни топологични пространства, $U \subseteq X$ и $F : U \to Y$. Казваме, че изображението $f : U \to \mathbb{R}$ има локален минимум в точката $\widetilde{x} \in \operatorname{int} U$ при условието $F(x) = 0$, ако $F(\widetilde{x}) = 0$ и съществува околност $V \subseteq U$ на $\widetilde{x}$ такава, че $f(x) \geq f(\widetilde{x})$ за всяко $x \in V$ такова, че $F(x) = 0$. Още казваме, че $\widetilde{x}$ е точка на *условен локален минимум* на $f(x)$ при условието $F(x) = 0$.

Аналогично се дефинира понятието за *условен локален максимум.* Общо условните локални минимуми и максимуми се наричат *условни локални екстремуми.*

Задачите за намиране на условен локален минимум ще записваме накратко по следния начин

$$\left| \begin{array}{l} f(x) \longrightarrow \inf \\ F(x) = 0. \end{array} \right. \tag{4.1}$$

При предположението, че и двете изображения f и F са диференцируеми по Фреше в своята дефиниционна област, такива задачи се наричат *гладки.*

Със задачата (4.1) се свързва изображението

$$\mathcal{L}(x, \lambda, y^*) := \lambda f(x) + \langle y^*, F(x) \rangle,$$

където $\lambda \in \mathbb{R}$ и $y^* \in Y^*$. Изображението $\mathcal{L}$ се нарича *функция на Лагранж*, а $\lambda \in \mathbb{R}$ и $y^* \in Y^*$ — *множители на Лагранж.*

В сила е следното необходимо условие за условен локален екстремум.

Теорема 4.6 (Правило на Лагранж)**.** *Нека X и Y са банахови пространства, $U \subseteq X$ е отворено и $\widetilde{x} \in U$. Нека още:*

а) изображението $f : U \to \mathbb{R}$ е диференцируемо по Фреше в точката $\widetilde{x}$;

б) изображението $F : U \to Y$ е диференцируемо по Фреше в околност на точката $\widetilde{x}$, $F'(x)$ е непрекъсната в $\widetilde{x}$ и $\operatorname{Im} F'(\widetilde{x})$ е затворено.

Тогава, ако $\widetilde{x}$ е точка на условен локален минимум в задачата (4.1), *то съществуват множители на Лагранж $\lambda \in \mathbb{R}$ и $y^* \in Y^*$, не и двата равни на нула, такива, че*

$$\mathcal{L}'_x(\widetilde{x}, \lambda, y^*) = \lambda f'(\widetilde{x}) + F'(\widetilde{x})^* y^* = 0. \tag{4.2}$$

Ако още $\operatorname{Im} F'(\widetilde{x}) = Y$, то (4.2) *е в сила с $\lambda = 1$.*

Преди да докажем теоремата ще направим няколко разяснения и ще изведем едно нейно важно следствие.

Съотношението (4.2) се нарича *уравнение на Ойлер-Лагранж на задачата* (4.1). В него с $\mathcal{L}'_x(\widetilde{x}, \lambda, y^*)$ сме означили производната на Фреше по x на функцията на Лагранж в точката $\widetilde{x}$; λ и y^* са фиксирани. Изображението $F'(\widetilde{x})^*$ е спрегнатият оператор на $F'(\widetilde{x})$. Да отбележим още, че $F'(\widetilde{x})^* y^*$ е производната на Фреше по x на изображението $\langle y^*, F(x)\rangle$ в точката $\widetilde{x}$. Това може да се види по следния начин. Първо да отбележим, че $\langle y^*, F(x)\rangle$ е функционал, дефиниран в X. След като $F(x)$ е диференцируемо по Фреше в $\widetilde{x}$, то

$$F(\widetilde{x} + h) = F(\widetilde{x}) + F'(\widetilde{x})h + o(\|h\|_X).$$

Оттук и линейността на y^* следва

$$\begin{aligned}\langle y^*, F(\widetilde{x} + h)\rangle &= \langle y^*, F(\widetilde{x})\rangle + \langle y^*, F'(\widetilde{x})h\rangle + \langle y^*, o(\|h\|_X)\rangle \\ &= \langle y^*, F(\widetilde{x})\rangle + \langle F'(\widetilde{x})^* y^*, h\rangle + o(\|h\|_X),\end{aligned}$$

като в преобразуването на второто събираемо горе вдясно използваме дефиницията на спрегнат оператор, а за третото — ограничеността на линейния функционал y^*. Полученото представяне показва, че ограниченият линеен функционал $F'(\widetilde{x})^* y^* : X \to \mathbb{R}$ удовлетворява дефиницията за производна на Фреше на $\langle y^*, F(x)\rangle$ в точката $\widetilde{x}$.[1]

Бележка 4.7. Често по-удобно е да записваме уравнението на Ойлер-Лагранж (4.2) във вида

$$\lambda\langle f'(\widetilde{x}), h\rangle + \langle y^*, F'(\widetilde{x})h\rangle = 0 \quad \forall h \in X. \tag{4.3}$$

Да забележим, че условието $F(x) = 0$ може да се представи във вида $\mathcal{L}_{y^*}(x, \lambda, y^*) = 0$. Действително първият член в дефиницията на функцията на Лагранж не зависи от y^* и следователно неговата производна на Фреше по y^* съществува и е нула. А вторият член е линеен функционал на y^* и следователно, както знаем, е диференцируем по Фреше и производната съвпада със самия този функционал. По-точно да положим $\mathcal{F}(y^*) := \langle y^*, F(x)\rangle$, където $x \in X$ е фиксирано. Така $\mathcal{F} : Y^* \to \mathbb{R}$ и $F(x) \in (Y^*)^* = Y^{**}$. За по-голяма яснота можем да гледаме на $\mathcal{F}(y^*)$

[1] Добра практика при пресмятането на производни на изображения, особено преди да натрупаме опит, е постоянно да си даваме сметка за дефиниционната област и областта от стойности на разглежданите оператори. Това ни предпазва от наивни грешки. Например в разгледания току-що случай, ако за по-голяма яснота положим $\mathcal{F}(x) := \langle y^*, F(x)\rangle$, имаме $\mathcal{F} : X \to \mathbb{R}$. Следователно, ако $\mathcal{F}'(\widetilde{x})$ съществува, то $\mathcal{F}'(\widetilde{x}) : X \to \mathbb{R}$, при това трябва да е линеен оператор. Така нашите пресмятания трябва да дадат именно такова изображение. След като $F : X \to Y$, то и $F'(\widetilde{x}) : X \to Y$. Следователно $F'(\widetilde{x})^* : Y^* \to X^*$, което дава $F'(\widetilde{x})^* y^* \in X^*$, т.е. $F'(\widetilde{x})^* y^* : X \to \mathbb{R}$, както и видяхме, че трябва да очакваме.

във вида $\mathcal{F}(y^*) := \langle F(x), y^* \rangle$.[1] Сега от (2.3) следва, че $\mathcal{F}'(y^*) = F(x)$ за всяко $y^* \in Y^*$.

Така твърдението на теоремата за случая $\operatorname{Im} F'(\widetilde{x}) = Y$ се свежда то това, че ако $\widetilde{x}$ е решение на (4.1), то съществува множител на Лагранж $\widetilde{y}^* \in Y^*$, такъв, че

$$\mathcal{L}'_x(\widetilde{x}, 1, \widetilde{y}^*) = 0, \quad \mathcal{L}'_{y^*}(\widetilde{x}, 1, \widetilde{y}^*) = 0,$$

т.е. $(\widetilde{x}, \widetilde{y}^*)$ удовлетворява необходимото условие на Ферма за безусловен екстремум, но за функцията $\mathcal{L}(x, 1, y^*)$. Това е характерен пример за т.нар. Принцип на Лагранж в теорията на екстремалните задачи. Той гласи следното.

Принцип на Лагранж. *Екстремална задача за условен екстремум може да се сведе (изобщо или само що се касае до необходимо условие) до задача за безусловен екстремум относно модифицирана функция, в чиято дефиниция участват дадена функция и условията на задачата.*

Този принцип играе фундаментална роля в решаването на екстремални задачи, но трябва да бъде установяван отделно за всеки тип такива задачи. Теорема 4.6 установява валидността на Принципа на Лагранж за гладки задачи от типа (4.1).

Една последна бележка от общ характер. Теорема 4.6 дава действително полезно необходимо условие за условен екстремум, само ако съществува $\lambda \neq 0$ с посоченото в нея свойство. Това е непременно така, ако $\operatorname{Im} F'(\widetilde{x}) = Y$. Но и дори, ако това условие не е налице или е трудно за проверяване, ние пак може да решим уравнението на Ойлер-Лагранж (4.2) и ако то има решение $\widetilde{x}$, за което $\lambda \neq 0$, можем да проверим дали $\widetilde{x}$ е решение на гладката задача (4.1) и така да намерим поне някои от точките на условен локален минимум.

Ще установим едно важно следствие на Теорема 4.6. То касае случая $Y = \mathbb{R}^m$. Ако още $X = \mathbb{R}^n$, то се свежда до познатата от основните курсове по анализ теорема за условни екстремуми. Разглеждаме гладката задача

$$\left| \begin{aligned} & f_0(x) \longrightarrow \inf \\ & f_i(x) = 0, \quad i = 1, \dots, m, \end{aligned} \right. \tag{4.4}$$

където $f_i : X \to \mathbb{R}$, $i = 0, 1, \dots, m$. Функцията на Лагранж за тази задача има вида

$$\mathcal{L}(x, \lambda_0, \lambda_1, \dots, \lambda_m) := \sum_{i=0}^{m} \lambda_i f_i(x),$$

където $x \in X$ и $\lambda_i \in \mathbb{R}$, $i = 0, 1, \dots, m$.

[1] В тази симетричност се състои удобството на означението $\langle \circ, \circ \rangle$, което използваме за действието на линейните функционали.

Следствие 4.8. *Нека X е банахово пространство, $U \subseteq X$ е отворено и $\widetilde{x} \in U$. Нека още функционалите $f_0, f_1, \dots, f_m$ са непрекъснато диференцируеми по Фреше в околност на точката $\widetilde{x}$. Тогава съществуват реални числа $\lambda_0, \lambda_1, \dots, \lambda_m$, не всички равни на нула, такива, че*

$$\sum_{i=0}^{m} \lambda_i f_i'(\widetilde{x}) = 0. \tag{4.5}$$

Ако функционалите $f_1'(\widetilde{x}), \dots, f_m'(\widetilde{x})$ са линейно независими, то (4.5) е в сила с $\lambda_0 = 1$.

Бележка 4.9. В случая $X = \mathbb{R}^n$, $f_i'(\widetilde{x})$ представлява градиента на функцията $f_i : U \to \mathbb{R}$ в точката $\widetilde{x}$.

Доказателство на Следствие 4.8. Твърдението следва непосредствено от Теорема 4.6 с $F(x) := (f_1(x), \dots, f_m(x))$ и $Y := \mathbb{R}^m$. Използваме, че всяко подпространство на пространство с крайна размерност е затворено и забелязваме, че условието $\operatorname{Im} F'(\widetilde{x}) = \mathbb{R}^m$ се свежда именно до линейната независимост на $f_1'(\widetilde{x}), \dots, f_m'(\widetilde{x})$. Ясно е, че $F'(x) = (f_1'(x), \dots, f_m'(x))$. □

Преминаваме към доказателството на теоремата за множителите на Лагранж.

Доказателство на Теорема 4.6. Нека $\operatorname{Im} F'(\widetilde{x}) = Y$. Прилагаме теоремата на Люстерник в точката $x_0 = \widetilde{x}$. Тогава за допирателното пространство на множеството $M = \{x \in U : F(x) = 0\}$ в $\widetilde{x}$ имаме

$$TM(\widetilde{x}) = \operatorname{Ker} F'(\widetilde{x}).$$

Следователно за всяко $h \in \operatorname{Ker} F'(\widetilde{x})$ съществуват $\varepsilon > 0$ и $r : [-\varepsilon, \varepsilon] \to X$ такива, че $\|r(t)\|_X = o(t)$ и $x_h(t) := \widetilde{x} + th + r(t) \in M$ за всяко $t \in [-\varepsilon, \varepsilon]$.[1]

Щом $x_h(t) \in M$, то $F(x_h(t)) = 0$ за всяко $t \in [-\varepsilon, \varepsilon]$. Разглеждаме реално-значната функция $\varphi(t) := f(x_h(t))$ в интервала $[-\varepsilon, \varepsilon]$. Тя достига локален минимум в точката $t = 0$. Функцията $\varphi(t)$ е диференцируема в тази точка. Действително, след като $\|r(t)\|_X = o(t)$, то от дефиницията за производна на Фреше непосредствено следва, че $x_h(t)$ е диференцируема в $t = 0$, като $(x_h)'(0) = h$. Според направените предположения, изображението $f(x)$ е диференцируемо в точката $\widetilde{x}$ и от теоремата за

[1] Величини от вида на $x_h(t)$ се наричат *вариации* в точката $\widetilde{x}$. Извеждането на необходими условия за екстремум се състои в разглеждането на подходящи вариации. В доказателството на необходимото условие на Ойлер за задачата със закрепени краища и в доказателството на обобщената теорема на Ферма използвахме най-простия вид вариации $x_h(t) := \widetilde{x} + th$.

диференциране по Фреше на съставни изображения следва, че $\varphi(t)$ е диференцируема в $t = 0$ и

$$\varphi'(0) = \langle f'(x_h(0)), (x_h)'(0)\rangle = \langle f'(\widetilde{x}), h\rangle.$$

Като приложим класическата теорема на Ферма, получаваме

$$0 = \varphi'(0) = \langle f'(\widetilde{x}), h\rangle.$$

Елементът $h \in \operatorname{Ker} F'(\widetilde{x})$ бе произволно фиксиран. Следователно $f'(\widetilde{x}) \in (\operatorname{Ker} F'(\widetilde{x}))^{\perp}$. От Лемата за анулатора имаме

$$(\operatorname{Ker} F'(\widetilde{x}))^{\perp} = \operatorname{Im} F'(\widetilde{x})^{*}.$$

Следователно съществува $y^* \in Y^*$ такова, че

$$f'(\widetilde{x}) = -F'(\widetilde{x})^* y^*,$$

откъдето и получаваме

$$f'(\widetilde{x}) + F'(\widetilde{x})^* y^* = 0.$$

С това твърдението на теоремата в случая $\operatorname{Im} F'(\widetilde{x}) = Y$ е установено.

Нека сега $\operatorname{Im} F'(\widetilde{x})$ е същинско подпространство на Y. Тогава теоремата на Хан-Банах (вж. Следствие 1.3) влече, че съществува ненулев $y^* \in Y^*$ такъв, че $\langle y^*, y\rangle = 0$ за всяко $y \in \operatorname{Im} F'(\widetilde{x})$. Така установихме, че за всяко $x \in X$ имаме

$$0 = \langle y^*, F'(\widetilde{x})x\rangle = \langle F'(\widetilde{x})^* y^*, x\rangle.$$

Това означава, че $F'(\widetilde{x})^* y^* = 0$. С това (4.2) е установено с $\lambda = 0$. □

Накрая ще приведем един пример за приложението на теоремата в безкрайномерния случай (X е функционално пространство).

Пример 4.1. Ще изведем необходимото условие на Ойлер за задачата със закрепени краища (Теорема 3.3). Прилагаме Теорема 4.6 с $X := C^1[t_0, t_1]$ (с норма $\|\circ\|_{C^1}$), $Y := \mathbb{R}^2$ (да речем, с евклидовата норма),

$$f(x) := \int_{t_0}^{t_1} L(t, x(t), \dot{x}(t))\, dt$$

и $F(x) := (x(t_0) - x_0, x(t_1) - x_1)$. Предполагаме, че $L(t, x, y)$ е непрекъснато диференцируема в някакво отворено подмножество на $\mathbb{R}^3$, което съдържа точките $(t, \widetilde{x}(t), \dot{\widetilde{x}}(t))$, $t \in [t_0, t_1]$, където $\widetilde{x}$ е точка на слаб локален минимум за задачата.

За да се уверим, че функционалът $f(x)$ е диференцируем по Фреше и намерим неговата производна, ще приложим теоремата за диференциране на съставни изображения и резултата от точка (в.3) в § 2. Дефинираме изображението $D : C^1[t_0, t_1] \to C([t_0, t_1], \mathbb{R}^2)$, като полагаме

$D(x) := (x, \dot{x})$. Това изображение е линейно и непрекъснато (ограничено) и следователно диференцируемо, като $D'(x) = D$ за всяко $x \in C^1[t_0, t_1]$ (вж. (1.3)). Въвеждаме още изображението $H : C([t_0, t_1], \mathbb{R}^2) \to \mathbb{R}$, дефинирано чрез

$$H(x, y) := \int_{t_0}^{t_1} L(t, x(t), y(t))\, dt, \quad x, y \in C[t_0, t_1].$$

Както установихме в § 2 (в.3), то е диференцируемо и (вж. (2.11))

$$H'(x, y)(h, k) = \int_{t_0}^{t_1} [L'_x(t, x(t), y(t))\, h(t) + L'_y(t, x(t), y(t))\, k(t)]\, dt, \quad h, k \in C[t_0, t_1],$$

за всеки $x, y \in C[t_0, t_1]$. Въз основа на горните разглеждания и представянето на f във вида

$$f(x) = H(D(x)),$$

установяваме, че f е диференцируема и

$$\langle f'(x), h\rangle = H'(D(x))D'(x)h = \int_{t_0}^{t_1} [L'_x(t, x(t), \dot{x}(t))\, h(t) + L'_{\dot{x}}(t, x(t), \dot{x}(t))\, \dot{h}(t)]\, dt, \quad h \in C^1[t_0, t_1]. \tag{4.6}$$

Горе отново сме означили, както е прието за този тип задачи, третата променлива на L формално с $\dot{x}$ и така пишем $L'_{\dot{x}}$ вместо L'_y.

Що се касае до изображението $F(x)$, то е афинно и следователно е диференцируемо по Фреше и (вж. (2.3))

$$F'(x)h = (h(t_0), h(t_1)) \quad \forall h \in C^1[t_0, t_1].$$

Очевидно $F'(x)$ е непрекъсната в операторната норма ($F'(x)$ не зависи от x) и $\operatorname{Im} F'(x) = \mathbb{R}^2$.

Готови сме да приложим Теорема 4.6. Тя дава, че съществува ограничен линеен функционал y^* над $\mathbb{R}^2$ такъв, че

$$\langle f'(\widetilde{x}), h\rangle + \langle F'(\widetilde{x})^* y^*, h\rangle = 0 \quad \forall h \in C^1[t_0, t_1]. \tag{4.7}$$

Ограничените линейни функционали y^* над $\mathbb{R}^2$ имат вида

$$\langle y^*, y\rangle = (\lambda_0, \lambda_1) \cdot (y_0, y_1) = \lambda_0 y_0 + \lambda_1 y_1, \quad y = (y_0, y_1) \in \mathbb{R}^2,$$

с някакъв вектор $(\lambda_0, \lambda_1) \in \mathbb{R}^2$. Тогава

$$\langle F'(\widetilde{x})^* y^*, h\rangle = \langle y^*, F'(\widetilde{x})h\rangle = \lambda_0 h(t_0) + \lambda_1 h(t_1). \tag{4.8}$$

Като в (4.7) вземем предвид формулите (4.6) и (4.8) получаваме, че ако $\widetilde{x}$ е точка на слаб локален минимум за задачата със закрепени краища, то съществуват $\lambda_0, \lambda_1 \in \mathbb{R}$ такива, че

$$\int_{t_0}^{t_1} [L'_x(t, \widetilde{x}(t), \dot{\widetilde{x}}(t))\, h(t) + L'_{\dot{x}}(t, \widetilde{x}(t), \dot{\widetilde{x}}(t))\, \dot{h}(t)]\, dt \\ + \lambda_0 h(t_0) + \lambda_1 h(t_1) = 0 \quad \forall h \in C^1[t_0, t_1].$$

Във частност за $h \in C^1[t_0, t_1]$ такова, че $h(t_0) = h(t_1) = 0$, горното равенство дава (3.6), което на свой ред влече, че $\widetilde{x}$ удовлетворява диференциалното уравнение на Ойлер-Лагранж (3.2).

Задачи

1. Докажете Теорема 4.4.
2. Посредством теоремата за множителите на Лагранж изведете необходимото условие на Ойлер за слаб локален минимум за:
   (а) задачата със свободен край (вж. Зад. 1 в § 3),
   (б) задачата на Болца (вж. Зад. 2 в § 3).
3. Посредством обобщената теорема на Ферма изведете необходимото условие на Ойлер за слаб локален минимум в задачата на Майер (вж. Зад. 3 в § 3).
4. Намерете n неотрицателни числа такива, че сумата от техните q-ти степени е равна на 1, а сумата на p-тите им степени е максимално голяма. Предполага се, че $p, q \geq 1$. Като следствие, изведете неравенството

$$\left(\frac{1}{n}\sum_{i=1}^{n} |x_i|^p\right)^{1/p} \leq \left(\frac{1}{n}\sum_{i=1}^{n} |x_i|^q\right)^{1/q} \qquad \forall x_i \in \mathbb{R}, \quad 1 \leq p < q.$$

5 Задача на Лагранж. Уравнения на Ойлер-Лагранж

Ще разгледаме един клас екстремални задачи върху пространства от векторно-значни функции на една реална променлива. По-точно, ще работим с изображения $x : [t_0, t_1] \to \mathbb{R}^n$, където $t_0, t_1 \in \mathbb{R}$ и $n \in \mathbb{N}$. Такива изображения еднозначно се дефинират посредством своите координатни функции $x_i : [t_0, t_1] \to \mathbb{R}$, $i = 1, 2, \ldots, n$. Така имаме

$$x(t) := (x_1(t), x_2(t), \ldots, x_n(t)).$$

Да припомним, че с $C([t_0, t_1], \mathbb{R}^n)$ ще означаваме банаховото пространство от всички непрекъснати изображения $x : [t_0, t_1] \to \mathbb{R}^n$ с норма $\|x\|_C := \max_{t \in [t_0, t_1]} |x(t)|$, където $|x(t)| := \sqrt{x_1^2(t) + x_2^2(t) + \cdots + x_n^2(t)}$ е стандартната евклидова норма в $\mathbb{R}^n$. А с $C^k([t_0, t_1], \mathbb{R}^n)$ ще означаваме банаховото пространство от всички k пъти непрекъснато диференцируеми изображения $x : [t_0, t_1] \to \mathbb{R}^n$ с норма $\|x\|_{C^k} := \max_{j=0,\ldots,k} \|x^{(j)}\|_C$.

Ще разгледаме следната екстремална задача, спадаща към класа на т.нар. *задачи на Лагранж*:

$$(5.1) \qquad \left| \begin{array}{ll} \mathcal{F}(x, u) := \displaystyle\int_{t_0}^{t_1} f(t, x(t), u(t))\, dt \longrightarrow \inf & \\ \dot{x}(t) = \varphi(t, x(t), u(t)), \quad t \in [t_0, t_1], & (*) \\ b_0(x(t_0)) = 0, \quad b_1(x(t_1)) = 0, & (**) \end{array} \right.$$

където

$$\begin{gathered} x \in C^1([t_0, t_1], \mathbb{R}^n), \quad u \in C([t_0, t_1], \mathbb{R}^m), \\ f : U \to \mathbb{R}, \quad \varphi : U \to \mathbb{R}^n, \quad U \subseteq \mathbb{R}^{n+m+1}, \\ b_i : U_i \to \mathbb{R}^{s_i}, \quad U_i \subseteq \mathbb{R}^n, \quad i = 0, 1. \end{gathered}$$

Тук $\dot{x}(t) = (\dot{x}_1(t), \dot{x}_2(t), \ldots, \dot{x}_n(t))$. Да припомним, че аналогично дефинираме интеграл (риманов или лебегов) от вектор-функция на една реална променлива, като полагаме за $x(t) := (x_1(t), x_2(t), \ldots, x_n(t))$

$$\int_{t_0}^{t_1} x(t)\, dt := \left(\int_{t_0}^{t_1} x_1(t)\, dt, \ldots, \int_{t_0}^{t_1} x_n(t)\, dt \right).$$

Да припомним още, че стандартното скаларно произведение на $a, b \in \mathbb{R}^n$ означаваме с $a \cdot b$, т.е.

$$a \cdot b := a_1 b_1 + a_2 b_2 + \cdots + a_n b_n,$$

където $a := (a_1, a_2, \ldots, a_n)$ и $b := (b_1, b_2, \ldots, b_n)$.

Двойка функции (x, u) наричаме *допустима точка* за задачата (5.1), ако $x \in C^1([t_0, t_1], \mathbb{R}^n)$, $u \in C([t_0, t_1], \mathbb{R}^m)$ и те удовлетворяват условията $(*)$ и $(**)$. В такъв случай още казваме, че функциите x и u са допустими.

Дефиниция 5.1. Казваме, че допустимата точка $(\widetilde{x}, \widetilde{u})$ е точка на слаб локален минимум на задачата (5.1), ако съществува $\delta > 0$ такова, че $\mathcal{F}(\widetilde{x}, \widetilde{u}) \le \mathcal{F}(x, u)$ за всички допустими точки (x, u) такива, че $\|x - \widetilde{x}\|_{C^1} < \delta$ и $\|u - \widetilde{u}\|_C < \delta$.

За сравнение, локалният минимум се нарича *силен*, ако условието $\|x - \widetilde{x}\|_{C^1} < \delta$ се замени с по-слабото $\|x - \widetilde{x}\|_C < \delta$.

За $t \in [t_0, t_1]$, $x \in C^1([t_0, t_1], \mathbb{R}^n)$, $u \in C([t_0, t_1], \mathbb{R}^m)$, $\lambda \in \mathbb{R}$ и $p \in C^1([t_0, t_1], \mathbb{R}^n)$ дефинираме следната функция

$$L(t, x, \dot{x}, u, \lambda, p) := \lambda f(t, x, u) + p \cdot (\dot{x} - \varphi(t, x, u)).$$

Тя се нарича *лагранжиан* на задачата (5.1). Функционалът

$$\mathcal{L}(x, u, \lambda, p, \ell_0, \ell_1) := \int_{t_0}^{t_1} L(t, x(t), \dot{x}(t), u(t), \lambda, p(t))\, dt \\ + \ell_0 \cdot b_0(x(t_0)) + \ell_1 \cdot b_1(x(t_1)),$$

където $\ell_i \in \mathbb{R}^{s_i}$, $i = 0, 1$, се нарича *функция на Лагранж* на (5.1), а λ, p, ℓ_0 и ℓ_1 се наричат *множители на Лагранж.*

Теорема 5.2. *Нека $m, n, s_0, s_1 \in \mathbb{N}$. Нека още $f, \varphi \in C^1(U)$ и $b_i \in C^1(U_i)$, $i = 0, 1$, където $U \subseteq \mathbb{R}^{n+m+1}$ и $U_i \subseteq \mathbb{R}^n$, $i = 0, 1$, са отворени. Ако $(\widetilde{x}, \widetilde{u})$ е точка на слаб локален минимум на задачата* (5.1), *то съществуват множители на Лагранж $\lambda \in \mathbb{R}$, $p \in C^1([t_0, t_1], \mathbb{R}^n)$ и $\ell_i \in \mathbb{R}^{s_i}$, $i = 0, 1$, не всички равни на нула, такива, че*

$$\frac{d}{dt} L'_{\dot{x}}(t, \widetilde{x}(t), \dot{\widetilde{x}}(t), \widetilde{u}(t), \lambda, p(t)) \\ = L'_x(t, \widetilde{x}(t), \dot{\widetilde{x}}(t), \widetilde{u}(t), \lambda, p(t)), \quad t \in [t_0, t_1], \tag{5.2}$$

$$L'_{\dot{x}}(t_i, \widetilde{x}(t_i), \dot{\widetilde{x}}(t_i), \widetilde{u}(t_i), \lambda, p(t_i)) = (-1)^i b'_i(\widetilde{x}(t_i))^* \ell_i, \quad i = 0, 1, \tag{5.3}$$

$$L'_u(t, \widetilde{x}(t), \dot{\widetilde{x}}(t), \widetilde{u}(t), \lambda, p(t)) = 0, \quad t \in [t_0, t_1]. \tag{5.4}$$

Диференциалните уравнения (5.2) и (5.4) (с x и u вместо съответно $\widetilde{x}$ и $\widetilde{u}$) се наричат *уравнения на Ойлер-Лагранж* на задачата (5.1), а граничните условия (5.3) се наричат *условия за трансверзалност.*

С L'_x, $L'_{\dot{x}}$ и L'_u сме означили градиентите на L съответно по променливите x, $\dot{x}$ и u, като третираме $\dot{x}$ формално като независима променлива. Така лявата и дясната страна на (5.2) представляват вектор-функции на t със стойности в $\mathbb{R}^n$. Подобно лявата страна на (5.4) представлява вектор-функции на t със стойности в $\mathbb{R}^m$, а дясната е нулевият вектор в $\mathbb{R}^m$. Лявата и дясната страна на (5.3) са вектори (точки) в $\mathbb{R}^n$. Да означим с $b_{i,k}$, $k = 1, \dots, s_i$, координатните функции на b_i. Имаме, че

$$b'_i(\widetilde{x}(t_i)) = \left(\frac{\partial b_{i,k}}{\partial x_j}(\widetilde{x}(t_i)) \right)_{k,j=1}^{s_i, n}$$

и следователно

$$b_i'(\widetilde{x}(t_i))^* = b_i'(\widetilde{x}(t_i))^T = \left(\frac{\partial b_{i,k}}{\partial x_j}(\widetilde{x}(t_i))\right)_{j,k=1}^{n,s_i}.$$

Да забележим, че[1]

$$\begin{aligned}
&L_x'(t,x,\dot{x},u,\lambda,p) = \lambda f_x'(t,x,u) - \varphi_x'(t,x,u)^* p,\\
&L_{\dot{x}}'(t,x,\dot{x},u,\lambda,p) = p,\\
&L_u'(t,x,\dot{x},u,\lambda,p) = \lambda f_u'(t,x,u) - \varphi_u'(t,x,u)^* p.
\end{aligned}$$

За краткост тук и по-нататък, когато е подходящо, изпускаме аргумента t. Да припомним, че $\varphi_x'(t,x,u)$ е пълната производна (производната на Фреше) на изображението φ по променливата $x := (x_1, x_2, \ldots, x_n)$ в точката (t,x,u). Тя се представя посредством матрица $n \times n$, чиито компоненти са частните производни на координатните функции на φ по $x_1, x_2, \ldots, x_n$ в тази точка. Аналогично $\varphi_u'(t,x,u)$ е пълната производна (производната на Фреше) на изображението φ по променливата $u := (u_1, u_2, \ldots, u_m)$ в точката (t,x,u). Тя се представя посредством матрица $n \times m$, чиито компоненти са частните производни на координатните функции на φ по $u_1, u_2, \ldots, u_m$ в тази точка. Матриците $\varphi_x'(t,x,u)^*$ и $\varphi_u'(t,x,u)^*$ са транспонираните матрици съответно на $\varphi_x'(t,x,u)$ и $\varphi_u'(t,x,u)$.

От последните формули следва, че (5.2)-(5.4) еквивалентно се записват във вида

$$\dot{p}(t) = \lambda f_x'(t,\widetilde{x}(t),\widetilde{u}(t)) - \varphi_x'(t,\widetilde{x},\widetilde{u}(t))^* p(t), \quad t \in [t_0,t_1], \tag{5.5}$$

$$p(t_i) = (-1)^i b_i'(\widetilde{x}(t_i))^* \ell_i, \quad i = 0,1, \tag{5.6}$$

$$\varphi_u'(t,\widetilde{x}(t),\widetilde{u}(t))^* p(t) = \lambda f_u'(t,\widetilde{x}(t),\widetilde{u}(t)), \quad t \in [t_0,t_1]. \tag{5.7}$$

Доказателство на Теорема 5.2. Ще приложим теоремата за множителите на Лагранж от предходната глава (Теорема 4.6). Въвеждаме следните означения

$$\begin{aligned}
&Z := C^1([t_0,t_1],\mathbb{R}^n) \times C([t_0,t_1],\mathbb{R}^m), \quad Y = C([t_0,t_1],\mathbb{R}^n),\\
&z := (x,u), \quad \|z\|_Z := \max\{\|x\|_{C^1}, \|u\|_C\},\\
&g(z) := \mathcal{F}(x,u), \quad g : Z \to \mathbb{R},\\
&F(z) := \dot{x} - \varphi(t,x,u), \quad F : Z \to Y,\\
&B_i(z) := b_i(x(t_i)), \quad B_i : Z \to \mathbb{R}^{s_i}, \quad i = 0,1,\\
&G(z) := (F(z), B_0(z), B_1(z)), \quad G : Z \to Y \times \mathbb{R}^{s_0} \times \mathbb{R}^{s_1}.
\end{aligned}$$

Нормата в $Y \times \mathbb{R}^{s_0} \times \mathbb{R}^{s_1}$ се въвежда за $(y,a,b) \in Y \times \mathbb{R}^{s_0} \times \mathbb{R}^{s_1}$ чрез $\|(y,a,b)\|_{Y\times\mathbb{R}^{s_0}\times\mathbb{R}^{s_1}} := \max\{\|y\|_Y, |a|, |b|\}$.

[1] Все пак уверете се самостоятелно във верността на тези формули.

Така задачата (5.1) приема вида

$$\left| \begin{aligned} & g(z) \longrightarrow \inf \\ & G(z) = 0. \end{aligned} \right. \tag{5.8}$$

Да забележим, че $(\widetilde{x}, \widetilde{u})$ е точка на слаб локален минимум за задачата (5.1) точно тогава, когато $\widetilde{z}$ е точка на (условен) локален минимум на задачата (5.8).

Ще проверим, че тази задача удовлетворява предположенията на теоремата за множителите на Лагранж.

(а) Функционалът g е диференцируем по Фреше в точката $\widetilde{z} = (\widetilde{x}, \widetilde{u})$. Това следва от направените разглеждания във втората част на (в.3) в § 2 с z вместо x, като вземем предвид, че $r(\zeta) = o(\|\zeta\|_C)$, $\zeta \in Z$, влече $r(\zeta) = o(\|\zeta\|_Z)$. А според (2.11) (вж. също (2.13)) за производната на Фреше на g в точката $\widetilde{z}$ имаме

$$g'(\widetilde{z})\zeta = \int_{t_0}^{t_1} [f'_x(t, \widetilde{x}(t), \widetilde{u}(t)) \cdot h(t) + f'_u(t, \widetilde{x}(t), \widetilde{u}(t)) \cdot k(t)] \, dt, \qquad \zeta = (h, k) \in Z. \tag{5.9}$$

(б) За да установим, че изображението $G(z)$ удовлетворява условията на теоремата за множителите на Лагранж, е достатъчно (и необходимо) да покажем, че всеки един от компонентите му $F(z)$, $B_0(z)$ и $B_1(z)$ ги удовлетворява.

(в) Изображението $F(z)$ е диференцируемо по Фреше и $F'(z)$ е непрекъсната в $\widetilde{z}$ в операторната норма. За да установим това и намерим производната на $F(z)$, използваме, че

$$F(z) = D(z) - \Phi(z),$$

където сме положили

$$D(z) = D(x, u) := \dot{x}, \quad \Phi(z) = \Phi(x, u) := \varphi(t, x, u).$$

Изображението D е линейно и ограничено; следователно диференцируемо и за всяко z имаме

$$D'(z)\zeta = D(\zeta) = \dot{h}, \quad \zeta = (h, k) \in Z. \tag{5.10}$$

От независимостта на производната от z следва и нейната непрекъснатост в операторната норма.

Диференцируемостта на Φ следва от § 2 (в.2), като отново вземем предвид твърдението, формулирано по-горе в (б), но отнесено към координатните функции на Φ. Още (2.10) с z вместо x дава

$$\Phi'(z)\zeta = \varphi'_x(t, x, u) \, h + \varphi'_u(t, x, u) \, k, \quad \zeta = (h, k) \in Z. \tag{5.11}$$

Тук отново използваме отбелязанато в (а), че $r(\zeta) = o(\|\zeta\|_C)$, $\zeta \in Z$, влече $r(\zeta) = o(\|\zeta\|_Z)$.

Що се касае до непрекъснатостта на $\Phi'(z)$ в операторната норма, тя следва от равномерната непрекъснатост на частните производни на координатните функции на φ върху всеки компакт в U.

За производната на $F(z)$ в точката $\widetilde{z}$ получаваме следната формула от (5.10) и (5.11)

$$F'(\widetilde{z})\zeta = \dot{h} - \varphi'_x(t, \widetilde{x}, \widetilde{u})\, h - \varphi'_u(t, \widetilde{x}, \widetilde{u})\, k, \quad \zeta = (h, k) \in Z. \tag{5.12}$$

(г) Изображенията $B_i(z)$, $i = 0, 1$, са диференцируеми по Фреше и $B_i'(z)$ са непрекъснати в $\widetilde{z}$ в операторната норма. Диференцируемостта, както и видът на производната бяха установени в § 2 (г). Според формула (2.12) имаме

$$B_i'(z)\zeta = b_i'(x(t_i))h(t_i), \quad \zeta = (h, k) \in Z, \quad i = 0, 1. \tag{5.13}$$

Аналогично на $\Phi'(z)$ се установява, че $B_i'(z)$ са непрекъснати в операторната норма.

(д) Ще покажем, че $\operatorname{Im} F'(\widetilde{z}) = Y$. Това означава, че за всяко $y \in Y$ съществува $z = (x, u) \in Z$ такова, че

$$F'(\widetilde{z})z = y.$$

Като вземем предвид (5.12), виждаме, че това е еквивалентно на твърдението, че системата линейни обикновени диференциални уравнения с неизвестни $x := (x_1, x_2, \ldots, x_n)$

$$\dot{x} - \varphi'_x(t, \widetilde{x}, \widetilde{u})\, x - \varphi'_u(t, \widetilde{x}, \widetilde{u})\, u = y$$

има решение в $C^1([t_0, t_1], \mathbb{R}^n)$. Но това, при направените предположения за непрекъснатост на коефициентите и на дясната част на уравненията, е добре известен факт (дори не само за някое, а за произволно $u \in C([t_0, t_1], \mathbb{R}^m)$).

(е) Множеството $\operatorname{Im} B_i'(\widetilde{z})$, $i = 0, 1$, е затворено в $\mathbb{R}^{s_i}$, защото е линейно подпространство на пространство с крайна размерност.

От (д) и (е) следва, че $\operatorname{Im} G'(\widetilde{z})$ е затворено в $Y \times \mathbb{R}^{s_0} \times \mathbb{R}^{s_1}$.

Така хипотезите на теоремата за множителите на Лагранж са проверени за задачата (5.8). Според нея щом $\widetilde{z}$ е точка на локален минимум за задачата (5.8), то съществуват $\lambda \in \mathbb{R}$ и ограничен линеен функционал ς^* над $Y \times \mathbb{R}^{s_0} \times \mathbb{R}^{s_1}$, не едновременно нула, такива, че

$$\widetilde{\mathcal{L}}'_z(\widetilde{z}, \lambda, \varsigma^*) = 0, \tag{5.14}$$

където $\widetilde{\mathcal{L}}$ е функцията на Лагранж за задачата (5.8), дефинирана чрез

$$\widetilde{\mathcal{L}}(z, \lambda, \varsigma^*) := \lambda g(z) + \langle \varsigma^*, G(z) \rangle.$$

За $\zeta \in Z$ имаме

$$\begin{aligned} \widetilde{\mathcal{L}}'_z(\widetilde{z}, \lambda, \varsigma^*)\zeta &= \langle \lambda g'(\widetilde{z}), \zeta \rangle + \langle G'(\widetilde{z})^* \varsigma^*, \zeta \rangle \\ &= \langle \lambda g'(\widetilde{z}), \zeta \rangle + \langle \varsigma^*, G'(\widetilde{z})\zeta \rangle \end{aligned} \tag{5.15}$$

Ще изведем представяне на $\langle \varsigma^*, G'(\widetilde{z})\zeta \rangle$ в термините на координатните компоненти на $G'(z) = (F'(z), B'_0(z), B'_1(z))$. От (1.1) и (1.5) следва, че всеки ограничен линеен функционал ς^* върху $Y \times \mathbb{R}^{s_0} \times \mathbb{R}^{s_1}$ се представя във вида

$$\langle \varsigma^*, w \rangle = \int_{t_0}^{t_1} y(t) \cdot d\mu(t) + \ell_0 \cdot a_0 + \ell_1 \cdot a_1, \quad w = (y, a_0, a_1) \in Y \times \mathbb{R}^{s_0} \times \mathbb{R}^{s_1},$$

където $\ell_i \in \mathbb{R}^{s_i}$, $i = 0, 1$, и $d\mu(t) = (d\mu_1(t), d\mu_2(t), \ldots, d\mu_n(t))$, като $\mu_i(t)$, $i = 1, \ldots, n$, са функции с ограничена вариация върху $[t_0, t_1]$, които са непрекъснати отляво в (t_0, t_1) и са равни на 0 в t_0.

От (5.14), предвид (5.15) и горната формула, следва, че съществуват $\lambda \in \mathbb{R}$, функции с ограничена вариация μ_i, $i = 1, \ldots, n$, върху $[t_0, t_1]$, и $\ell_i \in \mathbb{R}^{s_i}$, $i = 0, 1$, не едновременно нула, такива, че

$$\lambda g'(\widetilde{z})\zeta + \int_{t_0}^{t_1} [F'(\widetilde{z})\zeta](t) \cdot d\mu(t) + \ell_0 \cdot B'_0(\widetilde{z})\zeta + \ell_1 \cdot B'_1(\widetilde{z})\zeta = 0 \quad \forall \zeta \in Z. \tag{5.16}$$

Да въведем за краткост още следните означения:

$$\begin{aligned} &a(t) := f'_x(t, \widetilde{x}(t), \widetilde{u}(t)), \quad b(t) := f'_u(t, \widetilde{x}(t), \widetilde{u}(t)), \\ &A(t) := \varphi'_x(t, \widetilde{x}(t), \widetilde{u}(t)), \quad B(t) := \varphi'_u(t, \widetilde{x}(t), \widetilde{u}(t)), \\ &D_i := b'_i(\widetilde{x}(t_i)), \quad i = 0, 1. \end{aligned}$$

С тяхна помощ производните на Фреше в (5.9), (5.12) и (5.13) се представят във вида

$$\begin{aligned} &g'(\widetilde{z})\zeta = \int_{t_0}^{t_1} \big[a(t) \cdot h(t) + b(t) \cdot k(t)\big]\, dt, \\ &[F'(z)\zeta](t) = \dot{h}(t) - A(t)\, h(t) - B(t)\, k(t), \\ &B'_i(z)\zeta = D_i h(t_i), \quad i = 0, 1, \quad \zeta = (h, k) \in Z. \end{aligned} \tag{5.17}$$

В частност, от (5.16) с $\zeta = (h, 0)$, $h \in C^1([t_0, t_1], \mathbb{R}^n)$, и (5.17) получаваме

$$\begin{aligned} \int_{t_0}^{t_1} \lambda a(t) \cdot h(t)\, dt + \int_{t_0}^{t_1} \big[\dot{h}(t) - A(t)\, h(t)\big] \cdot d\mu(t) & \\ + \ell_0 \cdot D_0 h(t_0) + \ell_1 \cdot D_1 h(t_1) = 0.& \end{aligned} \tag{5.18}$$

Интегрираме по части интегралите[1] с $h(t)$:

$$\begin{aligned}\int_{t_0}^{t_1} \lambda a(t) \cdot h(t)\, dt &= \int_{t_0}^{t_1} h(t) \cdot \left(d \int_{t_1}^{t} \lambda a(\tau)\, d\tau\right) \\ &= h(t_0) \cdot \int_{t_0}^{t_1} \lambda a(\tau)\, d\tau + \int_{t_0}^{t_1} \left(\int_{t}^{t_1} \lambda a(\tau)\, d\tau\right) \cdot \dot{h}(t)\, dt.\end{aligned}$$

Аналогично постъпваме с втория интеграл. Тук изполваме формулата

$$\int_a^b \alpha(t)\gamma(t)\, d\mu(t) = \int_a^b \gamma(t)\, d\beta(t),$$

където $\alpha, \gamma \in C[a,b]$, $\mu \in BV[a,b]$ и

$$\beta(t) := -\int_t^b \alpha(\tau)\, d\mu(\tau), \quad t \in [a,b].$$

Да отбележим, че, при направените предположения, имаме $\beta \in BV[a,b]$. Тази формула може да се установи с теоремата на Радон-Никодим.

Така получаваме

$$\int_{t_0}^{t_1} A(t)\, h(t) \cdot d\mu(t) = h(t_0) \cdot \int_{t_0}^{t_1} A(\tau)^* d\mu(\tau) + \int_{t_0}^{t_1} \left(\int_{t}^{t_1} A(\tau)^*\, d\mu(\tau)\right) \cdot \dot{h}(t)\, dt.$$

$A(\tau)^*$ е всъщност транспонираната матрица на $A(\tau)$, изразът $A(\tau)^* d\mu$ е векторът, който се получава при умножаването на матрицата $A(\tau)^*$ и вектора $d\mu$, а интегралите по τ вдясно са интеграли от вектор-функции.

Преобразуваме и последното събираемо в (5.18) по следния начин:

$$\begin{aligned}\ell_1 \cdot D_1 h(t_1) &= (D_1^* \ell_1) \cdot h(t_1) \\ &= (D_1^* \ell_1) \cdot \left(h(t_0) + \int_{t_0}^{t_1} \dot{h}(\tau)\, d\tau\right) \\ &= (D_1^* \ell_1) \cdot h(t_0) + (D_1^* \ell_1) \cdot \int_{t_0}^{t_1} \dot{h}(\tau)\, d\tau \\ &= (D_1^* \ell_1) \cdot h(t_0) + \int_{t_0}^{t_1} (D_1^* \ell_1) \cdot \dot{h}(t)\, dt.\end{aligned}$$

Да положим

$$\begin{aligned}&(5.19) \qquad \nu(t) := \mu(t) + \int_{t_0}^{t} \left(\int_{s}^{t_1} \lambda a(\tau)\, d\tau - \int_{s}^{t_1} A(\tau)^* d\mu(\tau) + D_1^* \ell_1\right) ds, \\ &\qquad\qquad c := D_0^* \ell_0 + D_1^* \ell_1 + \int_{t_0}^{t_1} \lambda a(\tau)\, d\tau - \int_{t_0}^{t_1} A(\tau)^* d\mu(\tau).\end{aligned}$$

[1] Преобразуванията до края на доказателството се виждат непосредствено в едномерния случай. Общият изисква малко повече усилия и линейна алгебра. Полезно е самостоятелно да разгледате детайлите.

Да отбележим, че $\nu(t)$ е непрекъсната отляво в (t_0, t_1) и $\nu(t_0) = 0$.

Като заместим изведените току-що зависимости в (5.18), получаваме, че за всяко $h \in C^1([t_0, t_1], \mathbb{R}^n)$ е в сила

$$\int_{t_0}^{t_1} \dot{h}(t) \cdot d\nu(t) + c \cdot h(t_0) = 0. \tag{5.20}$$

Дясната страна на (5.20) представлява ограничен линеен функционал на h върху $C^1([t_0, t_1], \mathbb{R}^n)$. От единствеността на представянето на такива функционали в този вид следва, че

$$c = 0 \quad \text{и} \quad \nu_i(t) \equiv \text{const}, \; i = 1, \ldots, n.$$

Следователно $\mu_i \in AC[t_0, t_1]$, $i = 1, \ldots, n$, и ако положим $p_i := \dot{\mu}_i \in L[t_0, t_1]$, след диференциране на (5.19), получаваме

$$p(t) + \int_t^{t_1} \lambda a(\tau)\, d\tau - \int_t^{t_1} A(\tau)^* p(\tau)\, d\tau + D_1^* \ell_1 = 0 \quad \text{п.н. в } [t_0, t_1], \tag{5.21}$$

откъдето първо заключаваме, че можем да считаме, че $p_i \in AC[t_0, t_1]$ и (5.21) е в сила за всяко t, а след това, че $p_i \in C^1[t_0, t_1]$, $i = 1, \ldots, n$.

По-нататък, като положим $t = t_0$ в (5.21) и вземем предвид, че $c = 0$, получаваме

$$p(t_0) = D_0^* \ell_0,$$

а като положим $t = t_1$ в (5.21), получаваме

$$p(t_1) = -D_1^* \ell_1.$$

Така (5.6) е установено.

Като диференцираме (5.21) по t, получаваме

$$\dot{p}(t) = \lambda a(t) - A(t)^* p(t).$$

С това показахме (5.5).

Остана да докажем (5.7). За тази цел прилагаме (5.16) с $\zeta = (0, k)$, $k \in C([t_0, t_1], \mathbb{R}^m)$, и (5.17). Получаваме, че за всяко $k \in C([t_0, t_1], \mathbb{R}^m)$ е в сила

$$\begin{aligned} 0 &= \int_{t_0}^{t_1} \lambda b(t) \cdot k(t)\, dt - \int_{t_0}^{t_1} B(t)k(t) \cdot d\mu(t) \\ &= \int_{t_0}^{t_1} \lambda b(t) \cdot k(t)\, dt - \int_{t_0}^{t_1} B(t)k(t) \cdot p(t)\, dt \\ &= \int_{t_0}^{t_1} \lambda b(t) \cdot k(t)\, dt - \int_{t_0}^{t_1} B(t)^* p(t) \cdot k(t)\, dt \\ &= \int_{t_0}^{t_1} [\lambda b(t) - B(t)^* p(t)] \cdot k(t)\, dt. \end{aligned}$$

Това, както по-горе, влече

$$\lambda b(t) - B(t)^* p(t) \equiv 0,$$

с което и (5.7) е установено. □

Пример 5.1. Ще демонстрираме Теорема 5.2 отново върху задачата със закрепени краища

$$\left| \begin{array}{l} \displaystyle\int_{t_0}^{t_1} f(t, x(t), \dot{x}(t))\, dt \longrightarrow \inf \\ x(t_0) = x_0, \quad x(t_1) = x_1, \quad x \in C^1[t_0, t_1], \end{array} \right.$$

където сме предположили, че $f \in C^1(U)$, $U \subseteq \mathbb{R}^3$ е отворено. Тя се записва във вид на задача на Лагранж по следния начин:

$$\left| \begin{array}{l} \displaystyle\int_{t_0}^{t_1} f(t, x(t), u(t))\, dt \longrightarrow \inf \\ \dot{x}(t) = u(t), \quad t \in [t_0, t_1], \\ x(t_0) = x_0, \quad x(t_1) = x_1. \end{array} \right. \tag{5.22}$$

Прилагаме Теорема 5.2 с $n = m = 1$, $\varphi(t, x, u) := u$, $b_0(\xi) := \xi - x_0$ и $b_1(\xi) := \xi - x_1$, $\xi \in \mathbb{R}$. Имаме $b_i'(\xi) = 1$. Лагранжианът на задачата има вида

$$L(t, x, \dot{x}, u, \lambda, p) := \lambda f(t, x, u) + p(\dot{x} - u).$$

Теорема 5.2 влече, че ако $(\widetilde{x}, \widetilde{u})$ е точка на слаб локален минимум за задачата (5.22), то съществуват множители на Лагранж $\lambda \in \mathbb{R}$, $p \in C^1[t_0, t_1]$ и $\ell_1, \ell_2 \in \mathbb{R}$, не всичките равни на 0, такива, че (вж. (5.2)–(5.4))

$$\left| \begin{array}{l} \dfrac{d}{dt} L'_{\dot{x}} = L'_x \quad \Leftrightarrow \quad \dot{p}(t) = \lambda f'_x(t, \widetilde{x}(t), \widetilde{u}(t)), \\ L'_{\dot{x}}|_{t=t_i} = (-1)^i \ell_i \quad \Leftrightarrow \quad p(t_i) = (-1)^i \ell_i, \quad i = 0, 1, \\ L'_u = 0 \quad \Leftrightarrow \quad \lambda f'_u(t, \widetilde{x}(t), \widetilde{u}(t)) - p(t) = 0. \end{array} \right.$$

Ако допуснем, че $\lambda = 0$, то от последното уравнение горе получаваме, че $p(t) \equiv 0$. След това от уравненията на втория ред в системата горе получаваме $\ell_0 = \ell_1 = 0$. Така излиза, че всички множители на Лагранж са равни на 0, което е противоречие. Следователно $\lambda \neq 0$ и можем да считаме, че $\lambda = 1$.

Следователно

$$\frac{d}{dt} f'_u(t, \widetilde{x}(t), \widetilde{u}(t)) = \dot{p}(t) = f'_x(t, \widetilde{x}(t), \widetilde{u}(t)), \quad t \in [t_0, t_1].$$

Точката $(\widetilde{x}, \widetilde{u})$ е допустима и тогава $\widetilde{u} = \dot{\widetilde{x}}$. Следователно

$$\frac{d}{dt} f'_{\dot{x}}(t, \widetilde{x}(t), \dot{\widetilde{x}}(t)) = f'_x(t, \widetilde{x}(t), \dot{\widetilde{x}}(t)), \quad t \in [t_0, t_1].$$

Тук сме използвали уславянето частната производна на f по третата променлива u да я означаваме с $f'_{\dot{x}}$ вместо с f'_u. Последното уравнение е именно необходимото условие на Ойлер в Теорема 2.3.

Пример 5.2. Ще решим екстремалната задача

$$\int_0^1 \ddot{x}(t)^2\,dt \longrightarrow \inf,\ x(0)=\dot{x}(0)=0,\ x(1)=0,\ \dot{x}(1)=1, \tag{5.23}$$

в пространството $C^2[0,1]$. Представяме я във вид на задача на Лагранж по следния еквивалентен начин:

$$\left|\begin{array}{l} \displaystyle\int_0^1 u(t)^2\,dt \longrightarrow \inf \\ \dot{x}(t)=y(t), \quad t\in[0,1], \\ \dot{y}(t)=u(t), \quad t\in[0,1], \\ x(0)=y(0)=0, \\ x(1)=0,\ y(1)=1. \end{array}\right. \tag{5.24}$$

Пишем (x,y) вместо (x_1,x_2). Търсим решение в банаховото пространство $C^1([0,1],\mathbb{R}^2)$. Прилагаме Теорема 5.2 с $n=2$, $m=1$, $f(t,x,y,u):=u^2$, $\varphi(t,x,y,u):=(y,u)$, $b_0(\xi,\eta):=(\xi,\eta)$ и $b_1(\xi,\eta):=(\xi,\eta-1)$, $\xi,\eta\in\mathbb{R}$, $s_0=s_1=2$. Имаме, че $b'_i(\xi,\eta)$, $i=0,1$, представлява единичната квадратна матрица от ред 2. Лагранжианът на задачата има вида

$$L(t,x,y,\dot{x},\dot{y},u,\lambda,p_1,p_2):=\lambda u^2+p_1(\dot{x}-y)+p_2(\dot{y}-u).$$

Уравненията на Ойлер-Лагранж имат вида

$$\left|\begin{array}{l} \dfrac{d}{dt}L'_{\dot{x}}=L'_x \quad \Leftrightarrow \quad \dot{p}_1=0, \\ \dfrac{d}{dt}L'_{\dot{y}}=L'_y \quad \Leftrightarrow \quad \dot{p}_2=-p_1, \\ p_1(i)=(-1)^i\ell_{i,1}, \quad p_2(i)=(-1)^i\ell_{i,2}, \quad i=0,1, \\ L'_u=0 \quad \Leftrightarrow \quad 2\lambda u-p_2=0. \end{array}\right.$$

Отново, като в предходния пример, ако допуснем, че $\lambda=0$, получаваме, че всички останали множители на Лагранж са 0. Следователно $\lambda\neq 0$ и можем да считаме, че $\lambda=1$.

От системата горе веднага получаваме, че $u(t)$ е линейна функция. Но $u=\ddot{x}$. Следователно ако задачата има слабо решение $\widetilde{x}$, то непременно е кубичен полином. Единсвеният такъв полином, който удовлетворява граничните условия е $\widetilde{x}(t)=t^3-t^2$.

Остава да покажем, че $\widetilde{x}(t)$ е действително решение на задачата (5.23). За всяко $h\in C^2[0,1]$ такова, че $h(0)=\dot{h}(0)=h(1)=\dot{h}(1)=0$ имаме

$$\int_0^1\left[(\widetilde{x}(t)+h(t))^{\cdot\cdot}\right]^2dt=\int_0^1\ddot{\widetilde{x}}(t)^2\,dt+4\int_0^1(3t-1)\ddot{h}(t)\,dt+\int_0^1\ddot{h}(t)^2\,dt.$$

Третият интеграл горе вдясно е неотрицателен, а за втория получаваме, след като интегрираме по части и използваме граничните условия, които h удовлетворява, равенството

$$\int_0^1 (3t-1)\ddot{h}(t)\,dt = \int_0^1 (3t-1)\,d\dot{h}(t) = -3\int_0^1 \dot{h}(t)\,dt = 0.$$

Следователно

$$\int_0^1 \ddot{\widetilde{x}}(t)^2\,dt \le \int_0^1 \left[(\widetilde{x}(t)+h(t))^{\cdot\cdot}\right]^2 dt$$

за всяко $h \in C^2[0,1]$ такова, че $h(0) = \dot{h}(0) = h(1) = \dot{h}(1) = 0$.

С това установихме, че решението на дадената екстремална задача е $\widetilde{x}(t) = t^3 - t^2$. То е глобално.

Задачи

1. Уверете се във верността на твърдението, изказано в точка (б) от доказателството на Теорема 5.2.

2. Изведете от Теорема 5.2 необходимото условие на Ойлер за слаб локален минимум за задачата със свободни краища.

3. Изведете от Теорема 5.2 необходимото условие на Ойлер за слаб локален минимум за задачата на Болца.

4. Решете екстремалните задачи:

(а) $$\left|\begin{array}{l} \displaystyle\int_0^{\pi\sqrt{2}} (\ddot{x}^2 + x^2)\,dt \longrightarrow \inf, \\ x(0) = \dot{x}(0) = x(\pi\sqrt{2}) = 0, \\ \dot{x}(\pi\sqrt{2}) = -\operatorname{sh}\pi, \\ x \in C^2[0,\pi\sqrt{2}]; \end{array}\right.$$

(б) $$\left|\begin{array}{l} \displaystyle\int_0^1 (x^2 + u^2)\,dt \longrightarrow \inf, \\ \dot{x} = ax + u,\ x(0) = x_0,\ a \in \mathbb{R}, \\ x \in C^1[0,1],\ u \in C[0,1]; \end{array}\right.$$

(в) $$\left|\begin{array}{l} \displaystyle\int_0^1 (x^2 + u^2)\,dt \longrightarrow \inf \\ \ddot{x} + a\dot{x} = u,\ x(0) = x_0,\ a \in \mathbb{R} \\ x \in C^2[0,1],\ u \in C[0,1]. \end{array}\right.$$

6 Изпъкнали функции. Субдиференциали

Първо нека си припомним дефиницията и някои основни свойства на изпъкналите функции.

Дефиниция 6.1. Казваме, че едно подмножество D на линейно пространство е изпъкнало, ако D съдържа всяка отсечка, чиито краища са в него, т.е.

$$\alpha x_1 + (1-\alpha)x_2 \in D, \quad \forall x_1, x_2 \in D, \quad \forall \alpha \in [0,1].$$

Дефиниция 6.2. Казваме, че едно подмножество D на топологично линейно пространство е изпъкнало тяло, ако $\operatorname{int} D$ е изпъкнало и $D \subseteq \overline{\operatorname{int} D}$.

Тук $\operatorname{int} D$ обозначава вътрешността на множеството D, а $\overline{C}$ — затворената обвивка на множеството C.

Дефиниция 6.3. Нека D е изпъкнало множество. Казваме, че функцията $f : D \to \mathbb{R}$ е изпъкнала, ако

$$f(\alpha x_1 + (1-\alpha)x_2) \le \alpha f(x_1) + (1-\alpha) f(x_2) \quad \forall x_1, x_2 \in D, \quad \forall \alpha \in [0,1].$$

Дефиниция 6.4. Нека X е линейно пространство, $D \subseteq X$ и $f : D \to \mathbb{R}$. Надграфика на f наричаме множеството

$$\operatorname{epi}_D f := \{(x,y) \in X \times \mathbb{R} : x \in D,\ y \ge f(x)\}.$$

В случая $D = X$ ще пишем $\operatorname{epi} f$.

Разглеждаме линейното пространство $X \times \mathbb{R}$ със стандартната топология, породена от декартовото произведение на отворените подмножества на X и отворените интервали в $\mathbb{R}$.

В сила е следната основна теорема.

Теорема 6.5. *Нека D е изпъкнало тяло в топологично линейно пространство и $f : D \to \mathbb{R}$ е изпъкнала функция. Следните твърдения са еквивалентни:*

(а) f е ограничена отгоре в околност на вътрешната точка x_0;

(б) f е непркъсната във вътрешната точка x_0;

(в) f е непркъсната във $\operatorname{int} D$;

(г) $\operatorname{epi}_D f$ има вътрешни точки.

Доказателство. Ще докажем теоремата по схемата

$$(\text{а}) \implies (\text{б}) \implies (\text{в}) \implies (\text{г}) \implies (\text{а}).$$

(а) $\implies$ (б). Нека f е ограничена отгоре в околността $U \subseteq D$ на точката x_0, т.е. $f(x) \le M$ за $x \in U$ с някаква положителна константа M. Разглеждаме функцията $g(y) := f(y + x_0) - f(x_0)$ за $y \in U - x_0 := \{x - x_0 : x \in U\}$. Тя също е изпъкнала. Непрекъснатостта на $f(x)$ в точката x_0 е равносилна на непрекъснатостта на $g(y)$ в точката 0. За $0 < \varepsilon < M$ въвеждаме множеството

$$V_\varepsilon := \left(\frac{\varepsilon}{M}(U - x_0)\right) \cap \left(-\frac{\varepsilon}{M}(U - x_0)\right).$$

То е околност на 0. Ще покажем, че $|g(y)| \le \varepsilon$ за всяко $y \in V_\varepsilon$. След като $g(0) = 0$, това показва, че $g(y)$ е непрекъсната в точката 0. И така, нека $y \in (\varepsilon/M)(U - x_0)$, т.е. $(M/\varepsilon)y + x_0 \in U$. Тогава

$$g\left(\frac{M}{\varepsilon}\,y\right) = f\left(\frac{M}{\varepsilon}y + x_0\right) - f(x_0) =\le M$$

и от изпъкналостта на $g(y)$ следва, че

$$\begin{aligned} g(y) &= g\left(\frac{\varepsilon}{M}\,y + \left(1 - \frac{\varepsilon}{M}\right)0\right) \\ &\le \frac{\varepsilon}{M}\,g\left(\frac{M}{\varepsilon}\,y\right) + \left(1 - \frac{\varepsilon}{M}\right)g(0) \\ &\le \frac{\varepsilon}{M}\,M = \varepsilon. \end{aligned}$$

Аналогично, ако $y \in -(\varepsilon/M)(U - x_0)$, т.е. $-(M/\varepsilon)y + x_0 \in U$, то

$$g\left(-\frac{M}{\varepsilon}\,y\right) = f\left(-\frac{M}{\varepsilon}\,y + x_0\right) - f(x_0) \le M$$

и

$$\begin{aligned} 0 = g(0) &= g\left(\frac{1}{1+\varepsilon/M}\,y + \frac{\varepsilon/M}{1+\varepsilon/M}\left(-\frac{M}{\varepsilon}\,y\right)\right) \\ &\le \frac{1}{1+\varepsilon/M}\,g(y) + \frac{\varepsilon/M}{1+\varepsilon/M}\,g\left(-\frac{M}{\varepsilon}\,y\right) \\ &\le \frac{1}{1+\varepsilon/M}\,g(y) + \frac{\varepsilon/M}{1+\varepsilon/M}\,M \\ &= \frac{g(y)+\varepsilon}{1+\varepsilon/M}. \end{aligned}$$

Следователно

$$g(y) \ge -\varepsilon.$$

Така доказахме, че $|g(y)| \le \varepsilon$ за всяко $y \in V_\varepsilon$.

(б) $\Longrightarrow$ (в). Нека $x_1 \in \operatorname{int} D$ е произволна. Щом $f(x)$ е непрекъсната в x_0, то съществува околността $U \subseteq D$ на точката x_0 такава, че $|f(x) - f(x_0)| < 1$ за $x \in U$. Тогава $f(x) \le 1 + f(x_0)$ за $x \in U$. След като D е изпъкнало тяло, то съществува точка $x_2 \in D$ такава, че $x_1 = \alpha x_0 + (1-\alpha)x_2$ за някое $\alpha \in (0,1)$. Тогава $V := \alpha U + (1-\alpha)x_2$ е околност на x_1; при това тя се съдържа в $\operatorname{int} D$, защото $\operatorname{int} D$ е изпъкнало. Сега от изпъкналалостта пък на функцията $f(x)$ следва, че ако $y \in V$, т.е. $y = \alpha x + (1-\alpha)x_2$ за някое $x \in U$, то

$$\begin{aligned} f(y) &\le \alpha f(x) + (1-\alpha) f(x_2) \\ &\le \alpha(1 + f(x_0)) + (1-\alpha) f(x_2). \end{aligned}$$

Така установихме, че $f(x)$ е ограничена отгоре в околност на x_1 и тогава, според вече доказаното в предходния пункт, функцията е непрекъсната в тази точка.

(в) $\Longrightarrow$ (г). Тази импликация е тривиална. Нека $x_0 \in \operatorname{int} D$ и $U \subseteq D$ е нейна околност такава, че $f(x) \le 1 + f(x_0)$ за $x \in U$. Тогава множеството $U \times (f(x_0) + 1, \infty)$ е отворено подмножество на $\operatorname{epi}_D f$.

(г) $\Longrightarrow$ (а). Тук разсъжденията са също съвсем непосредствени. Нека $(x_0, y_0) \in \operatorname{int} \operatorname{epi}_D f$. Тогава (x_0, y_0) се съдържа с цялата своя околност $U \times (y_0 - \delta, y_0 + \delta)$ в $\operatorname{int} \operatorname{epi}_D f$; тук $\delta > 0$. Следователно $f(x) \le y_0 + \delta$ за всяко $x \in U$. $\square$

Непосредсвено от теоремата получаваме следното познато твърдение.

Следствие 6.6. *Всяка изпъкнала функция, дефинирана в изпъкнало тяло в $\mathbb{R}^n$, е непрекъсната във вътрешността на своята дефиниционна област.*

Преминаваме към основните свойства на изпъкналите функции по отношение на едно обобщение на понятието производна — т.нар. субградиент.

Дефиниция 6.7. Нека X е топологично линейно пространство и $f : X \to \mathbb{R}$. Казваме, че $x^* \in X^*$ е субградиент на f в точката $x_0 \in X$, ако

$$f(x) - f(x_0) \ge \langle x^*, x - x_0 \rangle \quad \forall x \in X.$$

Дефиниция 6.8. Нека X е топологично линейно пространство и $f : X \to \mathbb{R}$. Съвкупността от всички субградиенти на f в точката $x_0 \in X$ се нарича субдиференциал на f в x_0 и се означава чрез $\partial f(x_0)$.

Пример 6.1. Нека $X = \mathbb{R}$ и $f(x) := |x|$. Да припомним, че всеки ограничен линеен функционал x^* върху $\mathbb{R}$ се представя във вида

$$\langle x^*, x \rangle = c\,x, \quad x \in \mathbb{R},$$

за някое $c \in \mathbb{R}$.

Ако $x_0 > 0$, то $f(x) - f(x_0) = |x| - x_0$ и за да имаме

$$f(x) - f(x_0) \geq c(x - x_0) \quad \forall x \in \mathbb{R}, \tag{6.1}$$

трябва $c = 1$. Следователно $\partial f(x_0) = \{1\}$.

Аналогично се вижда, че ако $x_0 < 0$, то (6.1) е в сила само при $c = -1$ и следователно $\partial f(x_0) = \{-1\}$.

Нека $x_0 = 0$. Тогава

$$|x| \geq c\,x \quad \forall x \in \mathbb{R},$$

точно тогава, когато $c \in [-1, 1]$. Следователно $\partial f(0) = [-1, 1]$.[1]

Пример 6.2. Да разгледаме едно обобщение на функцията от предходния пример.

Нека X е нормирано линейно пространство и $f(x) := \|x\|$. Нека $x_0 \neq 0$. Тогава неравенството

$$\|x\| - \|x_0\| \geq \langle x^*, x - x_0 \rangle \quad \forall x \tag{6.2}$$

влече

$$\langle x^*, x - x_0 \rangle \leq \|x - x_0\| \quad \forall x,$$

откъдето $\|x^*\| \leq 1$.

От друга страна (6.2) с $x = 0$ влече

$$\langle x^*, x_0 \rangle \geq \|x_0\|.$$

Следователно

$$\|x^*\| = 1 \quad \text{и} \quad \langle x^*, x_0 \rangle = \|x_0\|. \tag{6.3}$$

Обратно, ако $x^* \in X^*$ удовлетворява (6.3), то

$$\langle x^*, x - x_0 \rangle = \langle x^*, x \rangle - \langle x^*, x_0 \rangle \leq \|x\| - \|x_0\| \quad \forall x.$$

Следователно x^* е субградиент на f в x_0.

Така установихме, че ако $x_0 \neq 0$, то

$$\partial \|x_0\| = \{x^* \in X^* : \|x^*\| = 1 \quad \text{и} \quad \langle x^*, x_0 \rangle = \|x_0\|\}.$$

Нека $x_0 = 0$. Условието, дефиниращо субградиентите на f в x_0, приема вида

$$\|x\| \geq \langle x^*, x \rangle \quad \forall x.$$

То е еквивалентно на $\|x^*\| \leq 1$. Следователно

$$\partial \|0\| = \{x^* \in X^* : \|x^*\| \leq 1\}.$$

[1]Използваме отъждествяването на ограничените линейни функционали върху $\mathbb{R}$ с $\mathbb{R}$ (вж. (1.1)). Можем да запишем резултата и без този факт. Нека означим с I идентитета върху $\mathbb{R}$. Тогава $\partial f(x_0) = \{I\}$ при $x_0 > 0$, $\partial f(x_0) = \{-I\}$ при $x_0 < 0$ и $\partial f(0) = \{cI : c \in [-1, 1]\}$.

Следващото твърдение съдържа няколко елементарни свойства на субдиференциала, които следват непосредсвено от дефиницията му.

Твърдение 6.9. *Нека X е топологично линейно пространство и $f : X \to \mathbb{R}$.*

(а) Ако $g(x) := f(x+a)$, $a \in X$, то $\partial g(x) = \partial f(x+a)$.

(б) Ако $g(x) := \lambda\, f(x)$, $\lambda > 0$, то $\partial g(x) = \lambda\, \partial f(x)$.

(в) Ако $g(x) := f(\lambda\, x)$, $\lambda > 0$, то $\partial g(x) = \lambda\, \partial f(\lambda\, x)$.

Лема 6.10. *Нека $\varphi : (a,b) \to \mathbb{R}$ е изпъкнала. Тогава φ има дясна производна $\varphi'_+(t)$ във всяка точка, при това*

$$\varphi'_+(t) = \inf_{\lambda>0} \frac{\varphi(t+\lambda)-\varphi(t)}{\lambda}.$$

Доказателство. Ще покажем, че за всяко фиксирано $t \in (a,b)$ диференчното частно $\frac{\varphi(t+\lambda)-\varphi(t)}{\lambda}$ е ограничена отдолу растяща функция на λ. Оттук ще следва, че то има граница при $\lambda \to 0+0$ и е налице формулата в лемата.

Понеже φ е изпъкнала, за всяко $\alpha \in (0,1)$ и достатъчно малко $\lambda > 0$ имаме

$$\varphi(t+\alpha\lambda) = \varphi(\alpha(t+\lambda)+(1-\alpha)t) \le \alpha\varphi(t+\lambda)+(1-\alpha)\varphi(t),$$

откъдето

$$\frac{\varphi(t+\alpha\lambda)-\varphi(t)}{\alpha\lambda} \le \frac{\varphi(t+\lambda)-\varphi(t)}{\lambda}.$$

Следователно $\frac{\varphi(t+\lambda)-\varphi(t)}{\lambda}$ е растяща функция на $\lambda > 0$.

Да фиксираме някое $t_0 \in (a,t)$. Аналогично получаваме

$$\begin{aligned}\varphi(t) &= \varphi\left(\frac{\lambda}{t+\lambda-t_0}\, t_0 + \frac{t-t_0}{t+\lambda-t_0}\,(t+\lambda)\right)\\ &\le \frac{\lambda}{t+\lambda-t_0}\,\varphi(t_0) + \frac{t-t_0}{t+\lambda-t_0}\,\varphi(t+\lambda),\end{aligned}$$

откъдето

$$\frac{\varphi(t+\lambda)-\varphi(t)}{\lambda} \ge \frac{\varphi(t)-\varphi(t_0)}{t-t_0}.$$

Така показахме, че $\frac{\varphi(t+\lambda)-\varphi(t)}{\lambda}$, като функция на λ, е ограничено отдолу. □

Твърдение 6.11. *Нека X е топологично линейно пространство и $f : X \to \mathbb{R}$ е изпъкнала. Тогава f има производна по всяка посока във всяка точка, при това*

$$f'(x;h) = \inf_{\lambda>0} \frac{f(x+\lambda h)-f(x)}{\lambda}.$$

Доказателство. Нека $x, h \in X$ са произволно фиксирани. От дефиницията на производна по посока имаме $f'(x; h) = \varphi'_+(0)$, където $\varphi(t) := f(x + th)$, $t \in (0, 1)$. Понеже f е изпъкнала, то и φ е такава. Сега твърдението следва от лемата. $\square$

Твърдение 6.12. *Ако изпъкнала функция, дефинирана върху топологично линейно пространство, е диференцируема по Гато в дадена точка, то производната ѝ на Гато в тази точка е единственият ѝ субградиент в нея.*

Доказателство. Нека X е топологично линейно пространство и $f : X \to \mathbb{R}$ е изпъкнала. Да предположим, че f е диференцируема по Гато в точката $x_0 \in X$. Тогава от дефиницията на производна на Гато и Твърдение 6.11 следва

$$\langle f'_G(x_0), h\rangle = f'(x_0; h) \leq f(x_0 + h) - f(x_0) \quad \forall h \in X.$$

Следователно $f'_G(x_0) \in \partial f(x_0)$.

Обратно нека $x^* \in \partial f(x_0)$, т.е.

$$f(x) - f(x_0) \geq \langle x^*, x - x_0\rangle \quad \forall x \in X.$$

Тогава

$$f(x_0 + th) - f(x_0) \geq \langle x^*, th\rangle = t\langle x^*, h\rangle \quad \forall t \in \mathbb{R} \quad \forall h \in X.$$

Следователно за произволно фиксирано $h \in X$ имаме

$$\frac{f(x_0 + th) - f(x_0)}{t} \geq \langle x^*, h\rangle, \quad t > 0.$$

След граничен преход $t \to 0 + 0$ в това неравенство, получаваме

$$\langle f'_G(x_0), h\rangle = f'(x_0; h) \geq \langle x^*, h\rangle.$$

Така установихме

$$\langle f'_G(x_0), h\rangle \geq \langle x^*, h\rangle \quad \forall h \in X.$$

Да забележим, че това неравенство с $-h$ на мястото на h, влече

$$\langle f'_G(x_0), h\rangle \leq \langle x^*, h\rangle \quad \forall h \in X.$$

От последните две неравенства сега получаваме

$$\langle f'_G(x_0), h\rangle = \langle x^*, h\rangle \quad \forall h \in X,$$

откъдето следва $x^* = f'_G(x_0)$. $\square$

Твърдение 6.13. *Всяка непрекъсната изпъкнала функция, дефинирана върху топологично линейно пространство, притежава субградиент във всяка точка.*

Доказателство. Нека X е топологично линейно пространство, $f : X \to \mathbb{R}$ е изпъкнала и непрекъсната и $x_0 \in X$. Прилагаме теоремата на Хан-Банах към множествата $A = \operatorname{int}(\operatorname{epi} f)$ и $B = \{(x_0, f(x_0))\}$. Те са непразни (вж. Теорема 6.5), непресичащи се и изпъкнали подмножества на топологичното линейно пространство $X \times \mathbb{R}$. Така установяваме, че съществува ненулев ограничен линеен функционал z^* над $X \times \mathbb{R}$ такъв, че

$$\langle z^*, a\rangle \le \langle z^*, (x_0, f(x_0))\rangle \quad \forall a \in A. \tag{6.4}$$

Ограниченият линеен функционал z^* се представя във вида

$$\langle z^*, a\rangle = \langle x^*, x\rangle + \alpha\, y, \quad a = (x, y) \in X \times \mathbb{R},$$

където $x^* \in X^*$ и $\alpha \in \mathbb{R}$. Така зависимостта в (6.4) приема вида

$$\langle x^*, x\rangle + \alpha\, y \le \langle x^*, x_0\rangle + \alpha\, f(x_0), \quad x \in X,\ y > f(x), \tag{6.5}$$

като x^* и α не са едновременно 0 понеже $z^* \neq 0$. Ако допуснем, че $\alpha = 0$, то ще получим от (6.5), че

$$\langle x^*, x - x_0\rangle \le 0 \quad \forall x \in X,$$

което влече $x^* = 0$. Следователно $\alpha \neq 0$.

Ако положим в (6.5) $x = x_0$, получаваме

$$\alpha(y - f(x_0)) \le 0 \quad \forall y > f(x_0).$$

Следователно $\alpha \le 0$.

Така установихме, че $\alpha < 0$. Разделяме двете страни на (6.5) на α, оставяме y да клони към $f(x)$ и полагаме $\overline{x}^* := -\alpha^{-1} x^*$. Така получаваме

$$f(x) - f(x_0) \ge \langle \overline{x}^*, x - x_0\rangle \quad \forall x \in X.$$

Следователно $\overline{x}^*$ е субградиент на f в точката x_0. □

Следва един силно нетривиален основен резултат, касаещ линейните свойства на субдиференциала на непрекъснатите изпъкнали функции.

Теорема 6.14 (Моро-Рокафелър)**.** *Нека X е топологично линейно пространство и $f_i : X \to \mathbb{R}$, $i = 1, 2, \ldots, n$, са изпъкнали функции. Тогава*

$$\partial(f_1 + f_2 + \cdots + f_n)(x) \supseteq \partial f_1(x) + \partial f_2(x) + \cdots + \partial f_n(x) \quad \forall x \in X.$$

Ако всички функции с изключение евентуално на една са непрекъснати, то

$$\partial(f_1 + f_2 + \cdots + f_n)(x) = \partial f_1(x) + \partial f_2(x) + \cdots + \partial f_n(x) \quad \forall x \in X.$$

Доказателство. Достатъчно е да установим теоремата за $n = 2$.

Първото твърдение следва непосредствено от дефиницията за субградиент. Действително нека $x_0 \in X$ е произволно фиксирано. Ако $\partial f_1(x_0) = \emptyset$ или $\partial f_2(x_0) = \emptyset$, то $\partial f_1(x_0) + \partial f_2(x_0) = \emptyset$ и твърдението е тривиално.

Нека $x_i^* \in \partial f_i(x_0)$, $i = 1, 2$. Тогава

$$f_i(x) - f_i(x_0) \geq \langle x_i^*, x - x_0 \rangle \quad \forall x \in X, \quad i = 1, 2.$$

Като „съберем“ тези две неравенства, получаваме

$$(f_1 + f_2)(x) - (f_1 + f_2)(x_0) \geq \langle x_1^* + x_2^*, x - x_0 \rangle \quad \forall x \in X,$$

откъдето $x_1^* + x_2^* \in \partial(f_1 + f_2)(x_0)$.

Преминаваме към доказателството на второто твърдение. Нека $x^* \in \partial(f_1 + f_2)(x_0)$ (ако $\partial(f_1 + f_2)(x_0) = \emptyset$, твърдението е тривиално). Нека $f_1(x)$ е непрекъсната в x_0. Трябва да покажем, че x^* се представя във вида $x^* = x_1^* + x_2^*$ за някои $x_1^* \in \partial f_1(x_0)$ и $x_2^* \in \partial f_2(x_0)$.

Тъй като $f_1(x)$ е непрекъсната в x_0, то, както следва от Теорема 6.5, $\operatorname{int}(\operatorname{epi} f_1) \neq \emptyset$. Разглеждаме следните подмножества на $X \times \mathbb{R}$:

$$A = \{(x, y) \in X \times \mathbb{R} : y \geq f_1(x + x_0) - f_1(x_0)\}$$

и

$$B = \{(x, y) \in X \times \mathbb{R} : y < \langle x^*, x \rangle - f_2(x + x_0) + f_2(x_0)\}.$$

Всъщност имаме $A = \operatorname{epi} f_1 - (x_0, f_1(x_0))$. Следователно A е изпъкнало и има непразна вътрешност. Както следва от изпъкналостта на f_2, B също е изпъкнало. Накрая, да отбележим, че $A \cap B = \emptyset$, защото иначе бихме получили, че съществува точка $(\overline{x}, \overline{y}) \in X \times \mathbb{R}$ такава, че

$$f_1(\overline{x} + x_0) - f_1(x_0) \leq \overline{y} < \langle x^*, \overline{x} \rangle - f_2(\overline{x} + x_0) + f_2(x_0).$$

Следователно

$$(f_1 + f_2)(\overline{x} + x_0) - (f_1 + f_2)(x_0) < \langle x^*, \overline{x} \rangle,$$

което противоречи на $x^* \in \partial(f_1 + f_2)(x_0)$.

Прилагаме теоремата на Хан-Банах към множествата A и B. Така получаваме, че съществуват $\widetilde{x}^* \in X^*$ и $\alpha \in \mathbb{R}$, не едновременно 0, такива, че

$$\langle \widetilde{x}^*, x_1 \rangle + \alpha\, y_1 \leq \langle \widetilde{x}^*, x_2 \rangle + \alpha\, y_2 \quad \forall (x_1, y_1) \in A \quad \forall (x_2, y_2) \in B. \tag{6.6}$$

След като y_1 не е ограничено отгоре, то $\alpha \leq 0$. Ако допуснем, че $\alpha = 0$, ще получим от (6.6), че

$$\langle \widetilde{x}^*, x_1 \rangle \leq \langle \widetilde{x}^*, x_2 \rangle \quad \forall x_1, x_2 \in X,$$

което влече $\widetilde{x}^* = 0$ — противоречие. Следователно $\alpha < 0$.

Разделяме двете страни на (6.6) на $-\alpha$, полагаме $x_1^* := -\alpha^{-1}\widetilde{x}^*$, взимаме $y_1 = f_1(x_1 + x_0) - f_1(x_0)$, и правим граничен преход $y_2 \to \langle x^*, x_2\rangle - f_2(x_2 + x_0) + f_2(x_0)$ отляво. Получаваме

$$\begin{aligned}\langle x_1^*, x_1\rangle &- f_1(x_1 + x_0) + f_1(x_0)\\ &\le \langle x_1^* - x^*, x_2\rangle + f_2(x_2 + x_0) - f_2(x_0) \quad \forall x_1, x_2 \in X.\end{aligned} \tag{6.7}$$

В частност за $x_2 = 0$ получаваме

$$\langle x_1^*, x_1\rangle \le f_1(x_1 + x_0) - f_1(x_0) \quad \forall x_1 \in X.$$

Следователно $x_1^* \in \partial f_1(x_0)$.

Накрая (6.7) с $x_1 = 0$ и $x_2^* := x^* - x_1^*$ дава

$$\langle x_2^*, x_2\rangle \le f_2(x_2 + x_0) - f_2(x_0) \quad \forall x_2 \in X,$$

което означава, че $x_2^* \in \partial f_2(x_0)$. □

Задачи

1. Посочете пример на изпъкнала функция, дефинирана в топологично линейно пространство, която не е непрекъсната в точката 0 (срв. с Теорема 6.5 и Следствие 6.6).

2. Опишете по-конкретно субдиференциала от Пример 6.2 за $X = L[a, b]$ и за $X = C[a, b]$.

3. Докажете, че ако функция, дефинирана върху топологично линейно пространство, има субградиент във всяка точка, то тя е изпъкнала (срв. с Твърдение 6.13).

4. Нека X е топологично линейно пространство, $f : X \to \mathbb{R}$ е изпъкнала и $x_0 \in X$. Тогава, както знаем (Твърдение 6.11) $f'(x_0, h)$ съществува по всяка посока $h \in X$. Докажете, че
$$x^* \in \partial f(x_0) \quad \iff \quad \langle x^*, h\rangle \le f'(x_0; h) \quad \forall h \in X.$$

5. Нека X е топологично линейно пространство и $f : X \to \mathbb{R}$ е непрекъсната и изпъкнала. Покажете, че
$$\sup_{x^* \in \partial f(x_0)} \langle x^*, x\rangle = f'(x_0; x), \quad x, x_0 \in X.$$

6. Покажете, че ако дадена непрекъсната изпъкнала функция, дефинирана върху топологично линейно пространство, притежава точно един субградиент в някоя точка, то тя е диференцируема по Гато в нея (срв. с Твърдение 6.12).

7 Изпъкнали задачи. Теорема на Кун-Такър

Нека X е топологично линейно пространство, $f_i : X \to \mathbb{R}$, $i = 0, 1, \ldots, n$, са изпъкнали функции и $C \subseteq X$ е изпъкнало множество. Ще разгледаме следната екстремална задача:

$$\tag{7.1} \left| \begin{array}{lr} f_0(x) \longrightarrow \inf & \\ f_i(x) \leq 0, \quad i = 1, 2, \ldots, n, & (*) \\ x \in C. & (**) \end{array} \right.$$

Задачи от този тип се наричат *изпъкнали.*

Бележка 7.1. Всяко локално решение на изпъкнала задача е и нейно глобално (абсолютно) решение. Действително, да предположим, че x_0 е локално решение на (7.1), което не е глобално, т.е. съществува $x_1 \neq x_0$, което удовлетворява условията $(*)$-$(**)$ и $f_0(x_1) < f_0(x_0)$. Разглеждаме точки от вида $x_\alpha := (1 - \alpha)x_0 + \alpha x_1$ за $\alpha \in (0, 1)$. От изпъкналостта на функциите f_i, $i = 1, \ldots, n$, и на множеството C следва, че x_α удовлетворява $(*)$-$(**)$ за всяко $\alpha \in (0, 1)$; уверете се в това самостоятелно. От изпъкналостта пък на f_0 следва, че

$$f_0(x_\alpha) \leq (1 - \alpha)f_0(x_0) + \alpha f_0(x_1) < f_0(x_0), \quad \alpha \in (0, 1).$$

Каквато и околност на точката x_0 да фиксираме, точката x_α ще попадне в нея стига да вземем α достатъчно малко положително число. Тук използваме свойството, че аритметичните операции са непрекъснати в топологичните линейни пространства. Така ще получим противоречие с това, че x_0 е локално решение на задачата (7.1).

Започваме с едно обобщение на теоремата на Ферма в случая на безусловен глобален минимум. То е формулирано в термините на субдиференциал. Това прави тази теорема особено естествена за непрекъснати изпъкнали функции, защото те именно (с точност непрекъснатост) притежават субградиент във всяка точка (вж. Твърдение 6.13 и Задача 3 в § 6) и за тях понятията за локален и глобален минимум съвпадат.

Теорема 7.2. *Нека X е топологично линейно пространство и $f : X \to \mathbb{R}$ е изпъкнала. Тогава точката $\widetilde{x} \in X$ е точка на глобален (абсолютен) минимум на $f(x)$ тогава и само тогава, когато $0 \in \partial f(\widetilde{x})$.*

Доказателство. Твърдението следва непосредствено от дефинициите.

Ако $\widetilde{x}$ е точка на абсолютен минимум на $f(x)$, то $f(x) - f(\widetilde{x}) \geq 0$ за всяко x. Следователно нулевият функционал върху X е субградиент на f в $\widetilde{x}$.

Обратно, ако нулевият функционал върху X е субградиент на f в $\widetilde{x}$, то за всяко $x \in X$ имаме $f(x) - f(\widetilde{x}) \geq \langle 0, x - \widetilde{x} \rangle = 0$. □

Пример 7.1. Една добра, макар и елементарна, илюстрация на тази теорема се дава от функцията $f(x) := \|x\|$ (вж. Пример 6.2).

Нека се върнем към общата задача, формулирана в началото. *Функцията на Лагранж* за задачата (7.1) има вида

$$\mathcal{L}(x, \lambda_0, \lambda_1, \ldots, \lambda_n) := \sum_{i=0}^{n} \lambda_i f_i(x),$$

където $\lambda_i \in \mathbb{R}$, $i = 0, 1, \ldots, n$, са *множителите на Лагранж.*

Следва основната характеризация на решенията на изпъкналите задачи.

Теорема 7.3 (Кун-Такър). *Нека X е топологично линейно пространство, $f_i : X \to \mathbb{R}$, $i = 0, 1, \ldots, n$, и $C \subseteq X$ са изпъкнали. Ако $\widetilde{x}$ е решение на задачата* (7.1)*, то съществуват множители на Лагранж $\lambda_0, \lambda_1, \ldots, \lambda_n$, които са неотрицателни, неедновременно* 0 *и такива, че*

$$\mathcal{L}(\widetilde{x}, \lambda_0, \lambda_1, \ldots, \lambda_n) = \min_{x \in C} \mathcal{L}(x, \lambda_0, \lambda_1, \ldots, \lambda_n) \tag{7.2}$$

и

$$\lambda_i f_i(\widetilde{x}) = 0, \quad i = 1, 2, \ldots, n. \tag{7.3}$$

Ако съществува $\xi \in C$ такова, че $f_i(\xi) < 0$, $i = 1, 2, \ldots, n$, то $\lambda_0 \neq 0$.

Обратно, ако точката $\widetilde{x}$ удовлетворява условията $(*)$ *и* $(**)$*, както и* (7.2)-(7.3) *с $\lambda_0 = 1$ и някакви $\lambda_i \geq 0$, $i = 1, 2, \ldots, n$, то тя е решение на задачата* (7.1).

Да отбележим, че ако $\lambda_0, \lambda_1, \ldots, \lambda_n$ са множители на Лагранж с посочените в теоремата свойства, то и $c\lambda_0, c\lambda_1, \ldots, c\lambda_n$, където $c > 0$, са също множители на Лагранж, които ги удовлетворяват. Затова, ако $\lambda_0 > 0$, то без ограничение можем да считаме, че $\lambda_0 = 1$.

Следва малко терминология:

- (7.2) се нарича *условие на Кун-Такър*,
- (7.3) се наричат *условия за допълняемост*,
- условието $\exists\, \xi \in C : f_i(\xi) < 0$, $i = 1, 2, \ldots, n$, се нарича *условие на Слейтър.*

Условието на Кун-Такър (7.2) показва, че Принципът на Лагранж е валиден в по-силна форма за изпъкналите задачи. По-точно, функцията на Лагранж достига своя абсолютен минимум при условието $(**)$ в решение на задача (7.1). Да забележим, че условието $(**)$ не е отразено в дефиницията на функцията на Лагранж.

Условията за допълняемост (7.3) показват, че всяко едно от условията $(*)$, което се изпълнява като строго неравенство в дадено решение на задачата, в действителност не участва във функцията на Лагранж. Казано иначе, ако предварително е известно, че дадена изпъкнала задача има единствено решение и то удовлетворява условие в групата $(*)$ като строго неравенство, това условие може да не се включва в дефиницията на функцията на Лагранж. Това лесно се разбира интуитивно, когато функцията, задаваща съответното условие, е непрекъсната (т.е. ограничена в околност на точката на условен екстремум). Щом даденото неравенство е строго в точката на екстремум, то от непрекъснатостта на тази функция следва, че то продължава да се изпълнява в цяла околност на точката на екстремум и следователно това условие не налага допълнителни ограничения. Можем да кажем, че такива условия не играят ефективна роля, те не са съществени и задачата не се променя, ако бъдат изпуснати.

При допълнителното предположение, че функциите са непрекъснати, Теорема 7.3 може да се формулира посредством субдиференциали. За тази цел ще въведем още две понятия.

Дефиниция 7.4. Нека $S \subseteq X$. *Индикаторна функция* на множеството S наричаме функцията

$$\delta_S(x) := \begin{cases} 0, & x \in S; \\ +\infty, & x \in X \setminus S. \end{cases}$$

Ако S е изпъкнало множество, то индикаторната му функция δ_S е изпъкнала.

Дефиниция 7.5. Нека X е топологично линейно пространство, $C \subseteq X$ изпъкнало множество и $x_0 \in C$. *Конус на опорните функционали* на C в точката x_0 наричаме множеството

$$N_C(x_0) := \{x^* \in X^* : \langle x^*, x - x_0 \rangle \le 0 \quad \forall x \in C\}.$$

Да забележим, че $N_C(x_0)$ е винаги непразно — нулевият функционал лежи във всеки такъв конус. Също така непосредствено се установява, че

$$N_C(x_0) = \partial \delta_C(x_0) \quad \text{за всяко} \quad x_0 \in C. \tag{7.4}$$

Сега можем да формулираме Теорема 7.3 още по следния начин стига участващите функции да са непрекъснати.

Теорема 7.6 (Кун-Такър, субдиференциална форма). *Нека X е топологично линейно пространство, $f_i : X \to \mathbb{R}$, $i = 0, 1, \ldots, n$, са изпъкнали и* ***непрекъснати****, и $C \subseteq X$ е изпъкнало. Ако $\widetilde{x}$ е решение на задачата*

(7.1)*, то съществуват множители на Лагранж* $\lambda_0, \lambda_1, \dots, \lambda_n$*, които са неотрицателни, неедновременно* 0 *и такива, че*

$$0 \in \lambda_0 \partial f_0(\widetilde{x}) + \lambda_1 \partial f_1(\widetilde{x}) + \cdots + \lambda_n \partial f_n(\widetilde{x}) + N_C(\widetilde{x}) \tag{7.5}$$

и

$$\lambda_i f_i(\widetilde{x}) = 0, \quad i = 1, 2, \dots, n. \tag{7.6}$$

Ако съществува $\xi \in C$ *такова, че* $f_i(\xi) < 0$*,* $i = 1, 2, \dots, n$*, то* $\lambda_0 \neq 0$*.*

Обратно, ако точката $\widetilde{x}$ *удовлетворява условията* $(*)$ *и* $(**)$*, както и* (7.5)-(7.6) *с* $\lambda_0 = 1$ *и някакви* $\lambda_i \geq 0$*,* $i = 1, 2, \dots, n$*, то тя е решение на задачата* (7.1)*.*

Условието (7.5) се нарича *уравнение на Ойлер-Лагранж.*

Бележка 7.7. Изпъкналостта на функцията позволява да се формулира не само необходимо, но и достатъчно условие. Сравнявайки двата типа предположения, които правим върху функцията, за да намерим екстремумите ѝ: диференцируемост или изпъкналост, би било погрешно да твърдим, че първото е по-силно, по-рестриктивно в сравнение с второто. Валидността на твърдение като Теорема 7.2 е едно доказателство за това. Друго идва от факта, че изпъкналостта влече диференцируемост, макар и в слаба форма — всяка изпъкнала функция на реална променлива има лява и дясна производна във всяка вътрешна точка, а в общия случай има производна по произволна посока.

Преминаваме към доказателството на двете теореми. Ще изведем втората от първата, а доказателството на първата е може би едно от най-красивите в курса.

Доказателство на Теорема 7.3. Нека A е множеството на всички точки $(\mu_0, \mu_1, \dots, \mu_n) \in \mathbb{R}^{n+1}$, за които съществува $x \in C$ такова, че

$$f_0(x) - f_0(\widetilde{x}) < \mu_0, \quad f_i(x) \leq \mu_i, \quad i = 1, 2 \dots, n.$$

Както следва от изпъкналостта на множестовото C и на функциите f_i, $i = 0, 1, \dots, n$, множеството A е изпъкнало (проверете това самостоятелно). Накрая да забележим, че $\mathbf{0} \notin A$, където $\mathbf{0} := (0, 0, \dots, 0) \in \mathbb{R}^{n+1}$, защото иначе бихме получили, че съществува $\overline{x} \in C$ такова, че $f_0(\overline{x}) < f_0(\widetilde{x})$ и $f_i(\overline{x}) \leq 0$, $i = 1, 2, \dots, n$, което противоречи на това, че $\widetilde{x}$ е решение на задачата (7.1) (вж. Бележка 7.1).

Прилагаме Теорема 1.6 към подмножествата A и $B := \{\mathbf{0}\}$ на $\mathbb{R}^{n+1}$.[1] Така получаваме, че съществува ненулев ограничен линеен функционал

[1] Вместо Теорема 1.6 можем да приложим теоремата на Хан-Банах за линейни топологични пространства (Теорема 1.5), но трябва допълнително да се убедим, че $\operatorname{int} A \neq \emptyset$. Последното следва от факта, че $A \supset \{(\mu_0, \mu_1, \dots, \mu_n) \in \mathbb{R}^{n+1} : \mu_i > 0 \ \forall i\}$.

z^* над $\mathbb{R}^{n+1}$ такъв, че

$$\langle z^*, \mu\rangle \leq \langle z^*, \mathbf{0}\rangle = 0 \quad \forall \mu := (\mu_0, \mu_1, \ldots, \mu_n) \in A.$$

Това означава, че съществуват реални числа $\lambda_0, \lambda_1, \ldots, \lambda_n$, не всички равни на 0 такива, че

$$\sum_{i=0}^{n} \lambda_i \mu_i \geq 0 \quad \forall (\mu_0, \mu_1, \ldots, \mu_n) \in A. \tag{7.7}$$

След като всички точки $(\mu_0, \mu_1, \ldots, \mu_n)$ с положителни координати са в A, то

$$\lambda_i \geq 0, \quad i = 0, 1, \ldots, n. \tag{7.8}$$

По-нататък да забележим, че за всеки $x \in C$ и $\varepsilon > 0$, точката с координати

$$\mu_0 = f_0(x) - f_0(\widetilde{x}) + \varepsilon, \quad \mu_i = f_i(x), \quad i = 1, 2, \ldots, n,$$

лежи в A. Тогава (7.7) с този избор на μ_i и след граничен преход $\varepsilon \to 0+0$ влече

$$\sum_{i=0}^{n} \lambda_i f_i(x) \geq \lambda_0 f_0(\widetilde{x}) \quad \forall x \in C. \tag{7.9}$$

Ако за някое $j \neq 0$ имаме $f_j(\widetilde{x}) = -\alpha < 0$, то A съдържа точката с координати $\mu_j = -\alpha$ и $\mu_i = \varepsilon$, $i \neq j$, за всяко $\varepsilon > 0$. Като приложим (7.7) с този избор на μ_i и направим граничен преход $\varepsilon \to 0 + 0$, получаваме $-\lambda_j \alpha \geq 0$. Следователно $\lambda_j \leq 0$, което предвид (7.8), дава $\lambda_j = 0$.

Така установихме, че ако $j \in \{1, 2, \ldots, n\}$, то

$$f_j(\widetilde{x}) < 0 \quad \Longrightarrow \quad \lambda_j = 0.$$

С това (7.3) е доказано.

Сега, като използваме (7.9) и (7.3), заключаваме, че

$$\sum_{i=0}^{n} \lambda_i f_i(x) \geq \sum_{i=0}^{n} \lambda_i f_i(\widetilde{x}) \quad \forall x \in C,$$

което дава (7.2).

Нека е изпълнено условието на Слейтър, т.е. нека съществува $\xi \in C$ такова, че $f_i(\xi) < 0$, $i = 1, 2, \ldots, n$. Ако допуснем, че $\lambda_0 = 0$, то, имайки предвид, че останалите λ_i са неотрицателни, като поне едно не е 0 и е в сила (7.3), установяваме, че

$$\sum_{i=0}^{n} \lambda_i f_i(\xi) < 0 = \sum_{i=0}^{n} \lambda_i f_i(\widetilde{x}),$$

което е в противоречие с (7.2). Следователно $\lambda_0 \neq 0$.

Преминаваме към доказателството на втората част на теоремата. Нека точката $\widetilde{x}$ удовлетворява условията $(*)$ и $(**)$, както и (7.2)-(7.3) с $\lambda_0 = 1$ и някакви $\lambda_i \geq 0$, $i = 1, 2, \dots, n$. Тогава за всяко x, удовлетворяващо $(*)$ и $(**)$, имаме

$$f_0(\widetilde{x}) \overset{(7.3)}{=} \sum_{i=0}^{n} \lambda_i f_i(\widetilde{x}) \overset{(7.2),\,(**)}{\leq} \sum_{i=0}^{n} \lambda_i f_i(x) \overset{\lambda_i \geq 0,\,(*)}{\leq} f_0(x).$$

Следователно $\widetilde{x}$ е решение на задачата (7.1). □

Доказателство на Теорема 7.6. Остава само да се докаже, че (7.2) и (7.5) са еквивалентни. Теорема 7.3 влече, че *разширената* функция на Лагранж

$$\mathcal{L}_1(x, \lambda_0, \lambda_1, \dots, \lambda_n) := \sum_{i=0}^{n} \lambda_i f_i(x) + \delta_C(x)$$

достига своя абсолютен минимум (относно x) в точката $\widetilde{x}$. Според Теорема 7.2 имаме

$$\mathcal{L}_1(\widetilde{x}, \lambda_0, \lambda_1, \dots, \lambda_n) = \min_{x \in X} \mathcal{L}_1(x, \lambda_0, \lambda_1, \dots, \lambda_n)$$

тогава и само тогава, когато

$$0 \in \partial_x \mathcal{L}_1(\widetilde{x}, \lambda_0, \lambda_1, \dots, \lambda_n).$$

Тук индексът x в ∂_x означава, че разглеждаме функцията $\mathcal{L}_1$ само като функция на x в дефиницията на нейния субдиференциал. Остава да приложим теоремата на Моро-Рокафелър (Теорема 6.14), Твърдение 6.9(б) и (7.4). □

Ще завършим с едно по-интересно приложение на Теорема 7.2

Пример 7.2. Нека са дадени три различни точки. Задачата е да се намери такава точка, че сумата от разстоянията между нея и дадените точки да бъде минимална.

Да означим дадените три различни точки с $y_1, y_2, y_3 \in \mathbb{R}^2$. Задачата се свежда към намирането на точката, в която изпъкналата функция

$$f(x) := |x - y_1| + |x - y_2| + |x - y_3|, \quad x \in \mathbb{R}^2.$$

достига своята най-малка стойност. Тук $|x|$ означава стандартната евклидова норма в равнината, т.е. дължината на отсечката между точката x и началото. Тази функция е непрекъсната и нараства неограничено с $|x|$ следователно достига своята най-малка стойност. Според Теорема 7.2

тя достига своята най-малка стойност в точката $\widetilde{x}$ тогава и само тогава, когато $0 \in \partial f(\widetilde{x})$. Теоремата на Моро-Рокафелър свежда последното към

$$0 \in \partial|\widetilde{x} - y_1| + \partial|\widetilde{x} - y_2| + \partial|\widetilde{x} - y_3|. \tag{7.10}$$

От резултата, който установихме в Пример 6.2 в частност за пространството $\mathbb{R}^2$ снабдено с нормата $|\cdot|$ имаме

$$\partial|x_0| = \begin{cases} \{\text{единичния вектор, еднопосочен с } x_0\}, & x_0 \neq \mathbf{0} \\ \{x \in \mathbb{R}^2 : |x| \leq 1\}, & x_0 = \mathbf{0}. \end{cases}$$

Следователно, ако $\widetilde{x} \neq y_1, y_2, y_3$, то (7.10) означава, че сумата от трите единични вектора, насочени от $\widetilde{x}$ съответно към точките y_1, y_2, y_3, е нула. Това на свой ред означава, че ъглите между тези вектори са равни на $120°$ (уверете се в това самостоятелно). Тогава ъглите на триъгълника с върхове y_1, y_2, y_3 са по-малки от $120°$. В този случай $\widetilde{x}$ се нарича точка на Торичели.

Нека $\widetilde{x}$ съвпада с някоя от дадените точки, да кажем y_1. Тогава от (7.10) следва, че дължината на сумата на единичните вектори, насочени от y_1 съответно към y_2 и y_3, не надминава единица. Оттук следва, че ъгълът при върха y_1 на триъгълника с върхове y_1, y_2, y_3 е поне $120°$ (отново проверете това самостоятелно).

И така, ако ъглите на триъгълника с върхове дадените точки са по-малки от $120°$, то решението на задачата е точката на Торичели, иначе търсената точка съвпада с върха на триъгълника, в който ъгълът е поне $120°$ (тук се включва и случаят, в който дадените точки са колинеарни).

Задачи

1. Докажете (7.4).
2. Намерете конуса на опорните функционали на
   (а) единичния кръг $x^2 + y^2 \leq 1$ в $\mathbb{R}^2$ в точката 0,
   (б) единичния кръг $x^2 + y^2 \leq 1,\ z = 0$ в $\mathbb{R}^3$ в точката 0,
   (в) единичния кръг $x^2 + y^2 \leq 1$ в $\mathbb{R}^2$ в точката $(1, 0)$,
   (г) конуса $|x| \leq y$ в $\mathbb{R}^2$ в точката $(0, 0)$.
3. Намерете решението на екстремалната задача, разгледана в Пример 7.2 при допълнителните условия:
   - търсената четвърта точка да принадлежи на дадено изпъкнало множество в равнината на дадените точки (разгледайте и конкретно случая, когато това множество е полуравнина);

- въведени са тегла в разстоянията до всяка една от дадените три точки.

4. Нека X е нормирано линейно пространство и $A_0, A_1, \dots, A_n$ са негови изпъкнали подмножества такива, че $A_0 \cap (\operatorname{int} A_1) \cap \cdots \cap (\operatorname{int} A_n) \neq \emptyset$. Да положим $A = A_0 \cap A_1 \cap \cdots \cap A_n$. Докажете, че

$$N_A(x) = N_{A_0}(x) + N_{A_1}(x) + \cdots + N_{A_n}(x), \quad x \in A.$$

8 Смесени задачи

Нека X и Y са банахови пространства, а U е множество. Нека още

$$\begin{aligned}&f_i : X \times U \to \mathbb{R}, \quad i = 0, 1, \ldots, n,\\&F : X \times U \to Y.\end{aligned}$$

Ще разгледаме екстермалната задача

$$(8.1) \qquad \left|\begin{array}{lr} f_0(x,u) \longrightarrow \inf & \\ F(x,u) = 0, & (*) \\ f_i(x,u) \le 0, \quad i = 1, 2, \ldots, n, & (**) \\ u \in U. & (***) \end{array}\right.$$

Задачи от този вид, стига функциите да удовлетворяват условията, които ще формулираме по-долу, се наричат *смесени*. Наименованието се дължи на това, че тези задачи представляват комбинация от гладка и изпъкнала задача. Смесените задачи са тяхно обобщение.

Целта ни ще бъде да установим необходими условия аз локален минимум.

Дефиниция 8.1. Ще казваме, че точката $(\widetilde{x}, \widetilde{u})$, удовлетворяваща условията $(*)$-$(***)$, е точка на локален минимум за задачата (8.1), ако

$$f_0(\widetilde{x}, \widetilde{u}) \le f_0(x, u)$$

за всички точки (x, u), които удовлетворяват $(*)$-$(***)$, x е в някаква ололност на $\widetilde{x}$, а $u \in U$ е произволно.

Отново въвеждаме *функция на Лагранж*. Тя влючва само функционалните условия на екстремалната задача и се дефинира посредством

$$\mathcal{L}(x, u, \lambda_0, \lambda_1, \ldots, \lambda_n, y^*) := \sum_{i=0}^{n} \lambda_i f_i(x, u) + \langle y^*, F(x, u) \rangle,$$

където $\lambda_i \in \mathbb{R}$, $i = 0, 1, \ldots, n$, и $y^* \in Y^*$ са *множителите на Лагранж*.

Теорема 8.2. *Нека X и Y са банахови пространства, а U е множество. Нека съществува околност $V \subseteq X$ на точката $\widetilde{x}$ такава, че*

(а) за всяко фиксирано $u \in U$ изображенията $f_i(x, u)$, $i = 0, 1, \ldots, n$, и $F(x, u)$ са диференцируеми по Фреше по x във V и производните им на Фреше са непрекъснати в $\widetilde{x}$;

(б) за всяко фиксирано $x \in V$ изображенията $f_i(x, u)$, $i = 0, 1, \ldots, n$, и $F(x, u)$ удовлетворяват следниите условия за изпъкналост:

$$\begin{aligned}&\forall u_1, u_2 \in U \quad \forall \alpha \in [0, 1] \quad \exists\, u \in U :\\&\qquad F(x, u) = \alpha F(x, u_1) + (1 - \alpha) F(x, u_2),\\&\qquad f_i(x, u) \le \alpha f_i(x, u_1) + (1 - \alpha) f_i(x, u_2), \quad i = 0, 1, \ldots, n.\end{aligned}$$

Накрая нека

(в) $L_0 := \operatorname{Im} F'_x(\widetilde{x}, \widetilde{u})$ *има крайна коразмерност в* Y.[1]

Ако $(\widetilde{x}, \widetilde{u})$ *е точка на локален минимум за задачата* (8.1), *то съществуват множители на Лагранж* $\lambda_0, \lambda_1 \dots, \lambda_n \geq 0$ *и* $y^* \in Y^*$, *не всички равни на* 0, *такива, че*

$$\mathcal{L}'_x(\widetilde{x}, \widetilde{u}, \lambda_0, \lambda_1, \dots, \lambda_n, y^*) = 0, \tag{8.2}$$

$$\mathcal{L}(\widetilde{x}, \widetilde{u}, \lambda_0, \lambda_1, \dots, \lambda_n, y^*) = \min_{u \in U} \mathcal{L}(\widetilde{x}, u, \lambda_0, \lambda_1, \dots, \lambda_n, y^*) \tag{8.3}$$

и

$$\lambda_i f_i(\widetilde{x}, \widetilde{u}) = 0, \quad i = 1, 2, \dots, n. \tag{8.4}$$

Както знаем,

$$\mathcal{L}'_x(\widetilde{x}, \widetilde{u}, \lambda_0, \lambda_1, \dots, \lambda_n, y^*) = \sum_{i=0}^{n} \lambda_i f'_{i_x}(\widetilde{x}, \widetilde{u}) + F'_x(\widetilde{x}, \widetilde{u})^* y^*.$$

Тук отново се използват следните наименования:

- (8.2) се нарича *уравнение на Ойлер-Лагранж*,
- (8.3) се нарича *условие на Кун-Такър*,
- (8.4) се наричат *условия за допълняемост*.

Условие (в) в теоремата следва от по-рестриктивното предположение

(в′) $\operatorname{Im} F'_x(\widetilde{x}, \widetilde{u}) = Y$.

Теорема 8.3. *В условията на Теорема 8.2, ако образът на множеството* $X \times U$ *под действието на изображението*

$$(x, u) \to F'_x(\widetilde{x}, \widetilde{u})x + F(\widetilde{x}, u)$$

съдържа околност на нулата в Y, *и ако съществува точка* $(\overline{x}, \overline{u})$ *такава, че*

$$F'_x(\widetilde{x}, \widetilde{u})\overline{x} + F(\widetilde{x}, \overline{u}) = 0,$$
$$\langle f'_{i_x}(\widetilde{x}, \widetilde{u}), \overline{x} \rangle + f_i(\widetilde{x}, \overline{u}) < 0$$

за всяко $i \in \{1, 2, \dots, n\}$, *за което* $f_i(\widetilde{x}, \widetilde{u}) = 0$, *то съществуват множители на Лагранж с посочните в Теорема 8.2 свойства, като при това още* $\lambda_0 = 1$.

[1] Това означава, че фактор-пространството Y/L_0 има крайна размерност.

Преминаваме към доказателствата на двете теореми.

Доказателство на Теорема 8.2. Нека $B := L_0 + F(\widetilde{x}, U)$ и L е линейната обвивка на B. От (б) следа, че множеството $F(\widetilde{x}, U)$ е изпъкнало. Множеството L_0 е също изпъкнало, защото е образ на линейно пространство под действието на линеен оператор. Следователно B е изпъкнало множество.

Известно е, че, ако образът на ограничен линеен оператор, действащ от едно банахово пространство в друго, има крайна коразмерност, то той е затворено множество.[1] Следователно L_0 и L са затворени подпространства на Y.

Разглеждаме няколко случая.

(1) Нека $L \neq Y$. От теоремата на Хан-Банах следва, че съществува ненулев ограничен линеен функционал y^* над Y такъв, че $\langle y^*, y \rangle = 0$ за всяко $y \in L$.[2] Така имаме

$$\langle y^*, F'_x(\widetilde{x}, \widetilde{u})x + F(\widetilde{x}, u) \rangle = 0 \quad \forall x \in X \quad \forall u \in U. \tag{8.5}$$

В частност за $u = \widetilde{u}$ имаме $F(\widetilde{x}, \widetilde{u}) = 0$ и следователно

$$\langle y^*, F'_x(\widetilde{x}, \widetilde{u})x \rangle = 0 \quad \forall x \in X.$$

След като

$$\langle y^*, F'_x(\widetilde{x}, \widetilde{u})x \rangle = \langle F'_x(\widetilde{x}, \widetilde{u})^* y^*, x \rangle,$$

получаваме

$$F'_x(\widetilde{x}, \widetilde{u})^* y^* = 0.$$

Аналогично, (8.5) с $x = 0$ влече

$$\langle y^*, F(\widetilde{x}, u) \rangle = 0 = \langle y^*, F(\widetilde{x}, \widetilde{u}) \rangle \quad \forall u \in U.$$

Следователно съотношения (8.2)-(8.4) са в сила с $\lambda_0 = \lambda_1 = \cdots = \lambda_n = 0$ и така фиксирания y^*. Тогава функцията на Лагранж има вида $\mathcal{L}(x, u, \lambda_0, \lambda_1, \ldots, \lambda_n, y^*) = \langle y^*, F(x, u) \rangle$.

(2) Нека $L = Y$. Ще покажем, че $\operatorname{int} B \neq \emptyset$. Да означим с $\pi : Y \to Y/L_0$ каноничното изображение на Y върху Y/L_0. Да припомним, че това е изображението, което съпоставя на елемент от Y класа на еквивалентност в Y/L_0, на който принадлежи. Така имаме, че $\pi y_1 = \pi y_2$ тогава и само тогава, когато $y_1 - y_2 \in L_0$.

[1] Това се показва с помощта на теоремата за отвореното изображение или затворената графика. Вижте напр. Proposition 7.10 в Klaus Deimling, Nonlinear Functional Analysis, Springer, 1985. Дължа благодарност на проф. Надежда Рибарска за този резултат.

[2] За да установим това, прилагаме споменатата теорема към L и някое кълбо в Y, което няма общи точки с L. Такова кълбо непременно съществува, защото L е затворено собствено подмножество на Y. Вижте Следствие 1.3.

Понеже линейната обвивка на B, т.е. L, съвпада с Y, то линейната обвивка на $\pi(B)$ съвпада с Y/L_0. Освен това понеже B е изпъкнало, то и $\pi(B)$ е такова. И така, $\pi(B)$ е изпъкнало множество, чиято линейна обвивка съвпада с цялото линейно пространство Y/L_0. Последното е с крайна размерност. Следователно $\operatorname{int}\pi(B) \neq \emptyset$.[1] Изображението π е непрекъснато. Следователно прообразът на всяко отворено подмножество на Y/L_0 (в частност и на отворено подмножество на $\pi(B)$) под действието на π е отворено подмножество на Y. Оттук, като вземем още предвид, че $\pi^{-1}(\pi(B)) = B$ (уверете се самостоятелно в последното) установяваме, че и $\operatorname{int} B \neq \emptyset$.

(2а) Ако $0 \notin \operatorname{int} B$, отново прилагаме теоремата на Хан-Банах този път към подмножествата B и $\{0\}$ на Y. Така установяваме, че съществува $y^* \in Y^*$ такъв, че

$$\langle y^*, y\rangle \geq 0 \quad \forall y \in B.$$

Това означава, че

$$\langle y^*, F'_x(\widetilde{x},\widetilde{u})x + F(\widetilde{x},u)\rangle \geq 0 \quad \forall x \in X \quad \forall u \in U. \tag{8.6}$$

Оттук, както в предния случай, полагайки $u = \widetilde{u}$, получаваме

$$\langle y^*, F'_x(\widetilde{x},\widetilde{u})x\rangle \geq 0 \quad \forall x \in X,$$

откъдето отново следва

$$F'_x(\widetilde{x},\widetilde{u})^* y^* = 0.$$

Пак аналогично на предния случай, полагаме $x = 0$ в (8.6). Така получаваме

$$\langle y^*, F(\widetilde{x},u)\rangle \geq 0 = \langle y^*, F(\widetilde{x},\widetilde{u})\rangle \quad \forall u \in U.$$

Така се убеждаваме, че отново множителите на Лагранж $\lambda_0 = \lambda_1 = \cdots = \lambda_n = 0$ и така фиксираният y^* притежават посочените в теоремата свойства.

(2б) Остана накрая да разгледаме случая $0 \in \operatorname{int} B$, като все така $L = Y$. Нека, евентуално след пренареждане на условията, имаме

$$f_1(\widetilde{x},\widetilde{u}) = \cdots = f_k(\widetilde{x},\widetilde{u}) = 0$$

и

$$f_{k+1}(\widetilde{x},\widetilde{u}) < 0, \ldots, f_n(\widetilde{x},\widetilde{u}) < 0.$$

[1] Твърдението, че всяко изпъкнало подмножество на $\mathbb{R}^d$, чиято линейна обвивка съвпада с $\mathbb{R}^d$, съдържа вътрешни точки, се свежда до твърдението, че стандартният симплекс в $\mathbb{R}^d$ има вътрешни точки.

Разглеждаме множеството A от онези $(\mu_0, \mu_1, \dots, \mu_k, y) \in \mathbb{R}^{k+1} \times Y$, за които съществуват $x \in X$ и $u \in U$ такива, че

$$\begin{aligned} &\mu_i > \langle f'_{i_x}(\widetilde{x}, \widetilde{u}), x\rangle + f_i(\widetilde{x}, u) - f_i(\widetilde{x}, \widetilde{u}), \quad i = 0, 1, \dots, k, \\ &y = F'_x(\widetilde{x}, \widetilde{u})x + F(\widetilde{x}, u) - F(\widetilde{x}, \widetilde{u}). \end{aligned}$$

Тук $F(\widetilde{x}, \widetilde{u}) = 0$ и $f_i(\widetilde{x}, \widetilde{u}) = 0$, $i = 1, 2, \dots, k$, но за следващите разсъждения е удобно да ги пишем.

Като използваме предположенията, направени в условие (б) на теоремата, установяваме, че A е изпъкнало (оставяме това за упражнение).

Ще покажем, че

$$\operatorname{int} A \neq \emptyset \quad \text{и} \quad \mathbf{0} \notin A, \tag{8.7}$$

където $\mathbf{0} = (0, 0, \dots, 0) \in \mathbb{R}^{k+1} \times Y$. Но преди това нека видим, че това е достатъчно, за да завършим доказателството на теоремата.

Прилагаме теоремата на Хан-Банах към множествата A и $\{\mathbf{0}\}$. Така установяваме, че съществува ненулев ограничен линеен функционал z^* над $\mathbb{R}^{k+1} \times Y$ такъв, че

$$\langle z^*, z\rangle \le \langle z^*, \mathbf{0}\rangle = 0 \quad \forall z := (\mu_0, \mu_1, \dots, \mu_k, y) \in A.$$

Това означава, че съществуват $\lambda_0, \lambda_1, \dots, \lambda_k \in \mathbb{R}$ и $y^* \in Y^*$, не всички равни на 0, такива, че

$$\sum_{i=0}^{k} \lambda_i \mu_i + \langle y^*, y\rangle \ge 0 \quad \forall (\mu_0, \mu_1, \dots, \mu_k, y) \in A. \tag{8.8}$$

Както в доказателството на теоремата на Кун-Такър, се вижда, че $\lambda_i \ge 0$, $i = 0, 1, \dots, k$, както и че (8.8) влече

$$\begin{aligned} &\sum_{i=0}^{k} \lambda_i \big[\langle f'_{i_x}(\widetilde{x}, \widetilde{u}), x\rangle + f_i(\widetilde{x}, u) - f_i(\widetilde{x}, \widetilde{u})\big] \\ &\qquad + \langle y^*, F'_x(\widetilde{x}, \widetilde{u})x + F(\widetilde{x}, u) - F(\widetilde{x}, \widetilde{u})\rangle \ge 0 \quad \forall x \in X \quad \forall u \in U, \end{aligned}$$

откъдето, като положим $\lambda_{k+1} = \dots = \lambda_n = 0$, получаваме

$$\begin{aligned} &\langle \mathcal{L}'_x(\widetilde{x}, \widetilde{u}, \lambda_0, \lambda_1, \dots, \lambda_n, y^*), x\rangle + \mathcal{L}(\widetilde{x}, u, \lambda_0, \lambda_1, \dots, \lambda_n, y^*) \\ &\qquad - \mathcal{L}(\widetilde{x}, \widetilde{u}, \lambda_0, \lambda_1, \dots, \lambda_n, y^*) \ge 0 \quad \forall x \in X \quad \forall u \in U. \end{aligned}$$

Аналогично на предните случаи, полагайки в последното съотношение веднъж $u = \widetilde{u}$, а след това $x = 0$, установяваме, че избраните множители на Лагранж удовлетворяват посочените в теоремата свойства.

Остава да се уверим във валидността на двете твърдения в (8.7).

За да установим, че $\operatorname{int} A \neq \emptyset$, можем да разсъждаваме по следния начин. От $0 \in \operatorname{int} B$ следва, че $0 \in \operatorname{int} \pi(B)$. Пространството Y/L_0 е крайномерно. Да означим с r неговата размерност. Тогава съществува r-мерен куб с център началото на Y/L_0, който се съдържа в $\pi(B)$. Да означим с $v_j \in U$, $j = 1, \ldots, m$, $m := r2^{r-1}$, прообразите на ръбовете на този куб под действието на изображението $\pi(F(\widetilde{x}, u))$.

Полагаме

$$O_1 := \{x \in X : \|x\| < 1\},$$

$$c := \max_{\substack{j=1,\ldots,m \\ i=0,\ldots,k}} \{f_i(\widetilde{x}, v_j) + \|f'_{i_x}(\widetilde{x}, \widetilde{u})\|\},$$

$$\begin{aligned} U_0 := \Bigg\{ u \in U : F(\widetilde{x}, u) = \sum_{j=1}^m \alpha_j F(\widetilde{x}, v_j), \\ f_i(\widetilde{x}, u) \leq \sum_{j=1}^m \alpha_j f_i(\widetilde{x}, v_j),\ i = 1, \ldots, k \\ \text{с някакви } \alpha_j \geq 0,\ \sum_{j=1}^m \alpha_j = 1 \Bigg\}, \end{aligned}$$

$$B_0 := F'_x(\widetilde{x}, \widetilde{u}) O_1 + F(\widetilde{x}, U_0).$$

От (б) следва, че B_0 е изпъкнало. Непосредствено се проверява, че $\pi(B_0)$ съдържа фиксирания по-горе куб. Тогава разглеждания съвсем аналогични на тези в началото на (2) показват, че $\operatorname{int} B_0 \neq \emptyset$. В този извод още се използва, че множеството $F'_x(\widetilde{x}, \widetilde{u}) O_1$ е отворено в L_0 — нещо, което следва от теоремата за отвореното изображение.

Сега разглеждаме множеството $A_0 := (\mathbb{R}_c)^{k+1} \times B_0$, където $\mathbb{R}_c := \{\mu \in \mathbb{R} : \mu > c\}$. Ясно е, че $A_0 \subseteq A$ и $\operatorname{int} A_0 \neq \emptyset$. С това първото твърдение в (8.7) е доказано.

Доказателството на $\mathbf{0} \notin A$ е доста технично. Нека първо накратко споменем каква е идеята. Допуска се противното. След това с помощта на теоремата на Люстерник и (б) се показва, че съществуват функции $x(t)$ и $u(t)$ със стойности съответно в X и U, дефинирани за достатъчно малки $t > 0$, такива, че $x = x(t)$ и $u = u(t)$ удовлетворяват условията $(*)$ и $(**)$ на екстремалната задача (8.1), $\|x(t) - \widetilde{x}\|_X \to 0$ при $t \to 0$, но $f_0(x(t), u(t)) < f_0(\widetilde{x}, \widetilde{u})$. Това означава, че $(\widetilde{x}, \widetilde{u})$ не е точка на локален минимум за (8.1), което е противоречие.

Преминаваме към подробното изложение на доказателството на $\mathbf{0} \notin A$. Допускаме противното, т.е съществуват $x_0 \in X$ и $u_0 \in U$ такива, че

$$F'_x(\widetilde{x}, \widetilde{u}) x_0 + F(\widetilde{x}, u_0) - F(\widetilde{x}, \widetilde{u}) = 0 \tag{8.9}$$

$$\langle f'_{i_x}(\widetilde{x}, \widetilde{u}), x_0 \rangle + f_i(\widetilde{x}, u_0) - f_i(\widetilde{x}, \widetilde{u}) \leq -\delta, \quad i = 0, 1, \ldots, k, \tag{8.10}$$

с някакво $\delta > 0$.

Фиксираме $\varepsilon > 0$ и разглеждаме изображението $G : X \times \mathbb{R}^{m+1} \to Y$, дефинирано посредством

$$G(x, \alpha_0, \dots, \alpha_m) := \left(1 - \alpha_0 - \varepsilon \sum_{j=1}^{m} \alpha_j\right) F(\widetilde{x} + x, \widetilde{u}) \\ + \alpha_0 F(\widetilde{x} + x, u_0) + \varepsilon \sum_{j=1}^{m} \alpha_j F(\widetilde{x} + x, v_j).$$

Ще приложим теоремата на Люстерник (Теорема 2.15) към него в точката $(0, 0, \dots, 0)$ в подходяща нейна околност.

Изображението G е диференцируемо по Фреше в множеството $(V - \widetilde{x}) \times \mathbb{R}^{m+1}$, G' е непрекъсната в нулата и

$$G'(0, 0, \dots, 0)(x, \alpha_0, \dots, \alpha_m) = F'_x(\widetilde{x}, \widetilde{u})\, x + \alpha_0 \big(F(\widetilde{x}, u_0) - F(\widetilde{x}, \widetilde{u})\big) \\ + \varepsilon \sum_{j=1}^{m} \alpha_j \big(F(\widetilde{x}, v_j) - F(\widetilde{x}, \widetilde{u})\big). \tag{8.11}$$

Ясно е, че $G(0, 0, \dots, 0) = F(\widetilde{x}, \widetilde{u}) = 0$. Също така непосредствено се вижда, че $\operatorname{Im} G'(0, 0, \dots, 0) \supseteq B_0$. Следователно $\operatorname{Im} G'(0, 0, \dots, 0) = Y$.

По-нататък понеже $\pi(F(\widetilde{x}, v_j))$, $j = 1, \dots, m$, са ръбовете на куб, то

$$\pi\Big(\sum_{j=1}^{m} F(\widetilde{x}, v_j)\Big) = 0,$$

т.е. $\sum_{j=1}^{m} F(\widetilde{x}, v_j) \in L_0$. Следователно съществува $x' \in X$ такова, че

$$F'_x(\widetilde{x}, \widetilde{u})x' + \sum_{j=1}^{m} F(\widetilde{x}, v_j) = 0. \tag{8.12}$$

От (8.9), (8.11) и (8.12) следва, че

$$G'(0, 0, \dots, 0)(x_0 + \varepsilon x', 1, \dots, 1) = 0,$$

откъдето, предвид $G(0, 0, \dots, 0) = 0$, получаваме $\|G(t(x_0 + \varepsilon x'), t, \dots, t)\|_Y = o(t)$.

Сега теоремата на Люстерник влече (уверете се в това самостоятелно), че съществуват $\varepsilon' > 0$ и изображения $\bar{x} : [-\varepsilon', \varepsilon'] \to X$ и $\bar{\alpha}_j : [-\varepsilon', \varepsilon'] \to \mathbb{R}$, $j = 0, \dots, m$, такива, че

$$G(t(x_0 + \varepsilon x' + \bar{x}(t)), t(1 + \bar{\alpha}_0(t)), \dots, t(1 + \bar{\alpha}_m(t))) = G(0, 0, \dots, 0) = 0$$

и $\|\bar{x}(t)\|_X \to 0$, $\bar{\alpha}_j(t) \to 0$, $j = 1, \dots, m$, при $t \to 0$.

Фиксираме ε така, че

$$\varepsilon\left|\langle f'_{i_x}(\widetilde{x},\widetilde{u}),x'\rangle+\sum_{j=1}^{m}f_i(\widetilde{x},u_j)\right|<\delta/2,\quad i=0,1,\dots,k. \tag{8.13}$$

Въвеждаме функциите

$$g_i(x,\alpha_0,\dots,\alpha_m):=\left(1-\alpha_0-\varepsilon\sum_{j=1}^{m}\alpha_j\right)f_i(\widetilde{x}+x,\widetilde{u})$$
$$+\alpha_0 f_i(\widetilde{x}+x,u_0)+\varepsilon\sum_{j=1}^{m}\alpha_j f_i(\widetilde{x}+x,v_j),\quad i=0,1,\dots,k.$$

За тях имаме

$$g_i(0,0,\dots,0)=\begin{cases}f_0(\widetilde{x},\widetilde{u}), & i=0,\\ 0, & i=1,\dots,k,\end{cases} \tag{8.14}$$

както и че са диференцируеми в точката $(0,0,\dots,0)$, като (срв. (8.11))

(8.15)
$$\langle g'_i(0,0,\dots,0),(x,\alpha_0,\dots,\alpha_m)\rangle=\langle f'_{i_x}(\widetilde{x},\widetilde{u}),x\rangle+\alpha_0\big(f_i(\widetilde{x},u_0)-f_i(\widetilde{x},\widetilde{u})\big)$$
$$+\varepsilon\sum_{j=1}^{m}\alpha_j\big(f_i(\widetilde{x},v_j)-f_i(\widetilde{x},\widetilde{u})\big).$$

Оттук и (8.10), предвид избора на ε (вж. (8.13)), следва

$$\langle g'_i(0,0,\dots,0),(x_0+\varepsilon x',1,\dots,1)\rangle<-\delta/2,\quad i=0,\dots,k.$$

От това, понеже

$$g_i(t(x_0+\varepsilon x'+\bar{x}(t)),t(1+\bar{\alpha}_0(t)),\dots,t(1+\bar{\alpha}_m(t)))=g_i(0,0,\dots,0)$$
$$+t\,\langle g'_i(0,0,\dots,0),(x_0+\varepsilon x',1,\dots,1)\rangle+o(t),$$

следва

$$g_i(t(x_0+\varepsilon x'+\bar{x}(t)),t(1+\bar{\alpha}_0(t)),\dots,t(1+\bar{\alpha}_m(t)))$$
$$\le g_i(0,0,\dots,0)-t\delta/2+o(t). \tag{8.16}$$

Полагаме $x(t):=\widetilde{x}+t(x_0+\varepsilon x'+\bar{x}(t))$. Тогава $\|x(t)-\widetilde{x}\|_X\to 0$ при $t\to 0$ и за достатъчно малки по абсолютна стойност t имаме $x(t)\in V$. Считаме още, че t е толкова близко до 0, че

$$t\left(1+\bar{\alpha}_0(t)+\varepsilon\sum_{j=1}^{m}\big(1+\bar{\alpha}_j(t)\big)\right)\le 1$$

и

$$1 + \bar{\alpha}_j(t) \geq 0, \quad j = 0, \ldots, m.$$

Да забележим, че от направеното в (б) предположение за две точки u_1 и u_2 по индукция се установява неговото обобщение за произволен краен брой точки. Така за всяко достатъчно малко по абсолютна стойност t съществува $u(t) \in U$ такова, че

$$\begin{aligned} F(x(t), u(t)) &= \left(1 - t\big(1 + \bar{\alpha}_0(t)\big) - \varepsilon t \sum_{j=1}^{m} \big(1 + \bar{\alpha}_j(t)\big)\right) F(x(t), \widetilde{u}) \\ &\quad + t\big(1 + \bar{\alpha}_0(t)\big) F(x(t), v_0) + \varepsilon t \sum_{j=1}^{m} \big(1 + \bar{\alpha}_j(t)\big) F(x(t), v_j) \\ &= G(t(x_0 + \varepsilon x' + \bar{x}(t)), t(1 + \bar{\alpha}_0(t)), \ldots, t(1 + \bar{\alpha}_m(t))) = 0, \end{aligned}$$

както и

$$\begin{aligned} f_i(x(t), u(t)) &\leq \left(1 - t\big(1 + \bar{\alpha}_0(t)\big) - \varepsilon t \sum_{j=1}^{m} \big(1 + \bar{\alpha}_j(t)\big)\right) f_i(x(t), \widetilde{u}) \\ &\quad + t\big(1 + \bar{\alpha}_0(t)\big) f_i(x(t), v_0) + \varepsilon t \sum_{j=1}^{m} \big(1 + \bar{\alpha}_j(t)\big) f_i(x(t), v_j) \\ &= g_i(t(x_0 + \varepsilon x' + \bar{x}(t)), t(1 + \bar{\alpha}_0(t)), \ldots, t(1 + \bar{\alpha}_m(t))), \\ &\qquad\qquad i = 0, \ldots, n. \end{aligned}$$

Оттук, (8.14) и (8.16) следва, че за малки $t > 0$ имаме

$$\begin{aligned} &F(x(t), u(t)) = 0, \\ &f_0(x(t), u(t)) < f_0(\widetilde{x}, \widetilde{u}), \\ &f_i(x(t), u(t)) < 0, \quad i = 1, \ldots, k, \end{aligned}$$

както и

$$\limsup_{t \to 0} f_i(x(t), u(t)) \leq f_i(\widetilde{x}, \widetilde{u}) < 0, \quad i = k+1, \ldots, n.$$

Да припомним още, че $\|x(t) - \widetilde{x}\|_X \to 0$ при $t \to 0$ и $u(t) \in U$.

Така излезе, че $(\widetilde{x}, \widetilde{u})$ не е локален минимум за задачата (8.1), което е противоречие. Това завършва доказателството на второто твърдение в (8.7), както и на теоремата. □

Доказателство на Теорема 8.3. Разсъжденията са подобни на тези, с които установихме аналогичното твърдение в теоремата на Кун-Такър.

Допускаме, че $\lambda_0 = 0$. Нека k е както в предното доказателство. Ако поне едно от $\lambda_1, \ldots, \lambda_k$ не е нула, т.е. е положително, получаваме, че

$$\sum_{i=1}^{k} \lambda_i \big[\langle f'_{i_x}(\widetilde{x}, \widetilde{u}), \overline{x}\rangle + f_i(\widetilde{x}, \overline{u})\big] < 0.$$

Сега, като използваме още (8.2) и (8.4), закючаваме, че

$$-\langle y^*, F'_x(\widetilde{x}, \widetilde{u})\overline{x}\rangle + \sum_{i=0}^{n} \lambda_i f_i(\widetilde{x}, \overline{u}) < 0.$$

Оттук, като вземем предвид съотношението

$$F'_x(\widetilde{x}, \widetilde{u})\overline{x} + F(\widetilde{x}, \overline{u}) = 0,$$

получаваме

$$\langle y^*, F(\widetilde{x}, \overline{u})\rangle + \sum_{i=0}^{n} \lambda_i f_i(\widetilde{x}, \overline{u}) < 0,$$

т.е.

$$\mathcal{L}(\widetilde{x}, \overline{u}, \lambda_0, \lambda_1, \ldots, \lambda_n, y^*) < 0 = \mathcal{L}(\widetilde{x}, \widetilde{u}, \lambda_0, \lambda_1, \ldots, \lambda_n, y^*),$$

което обаче е в противоречие с (8.3). Следователно $\lambda_1 = \cdots = \lambda_k = 0$.

Така получихме, че $\lambda_i = 0$ за всяко i, откъдето следва, че $y^* \neq 0$, а функцията на Лагранж придобива вида $\mathcal{L}(x, u, \lambda_0, \lambda_1, \ldots, \lambda_n, y^*) = \langle y^*, F(x, u)\rangle$.

Щом $y^* \neq 0$, то във всяка околност на нулата на Y, в частност в околността в условието на теоремата, съществува y_0 такова, че $\langle y^*, y_0\rangle < 0$. Нека $x_0 \in X$ и $u_0 \in U$ са такива, че

$$y_0 = F'_x(\widetilde{x}, \widetilde{u})x_0 + F(\widetilde{x}, u_0).$$

Следователно

$$\langle y^*, F'_x(\widetilde{x}, \widetilde{u})x_0 + F(\widetilde{x}, u_0)\rangle < 0. \tag{8.17}$$

Според (8.2) имаме $F'_x(\widetilde{x}, \widetilde{u})^* y^* = 0$ и следователно

$$\langle y^*, F'_x(\widetilde{x}, \widetilde{u})x_0\rangle = \langle F'_x(\widetilde{x}, \widetilde{u})^* y^*, x_0\rangle = 0.$$

Предвид на това (8.17) придобива вида

$$\langle y^*, F(\widetilde{x}, u_0)\rangle < 0,$$

което влече

$$\mathcal{L}(\widetilde{x}, u_0, \lambda_0, \lambda_1, \ldots, \lambda_n, y^*) < 0 = \mathcal{L}(\widetilde{x}, \widetilde{u}, \lambda_0, \lambda_1, \ldots, \lambda_n, y^*)$$

— отново противоречие с (8.3).

Така окончателно показахме, че допускането $\lambda_0 = 0$ води непременно до противоречие. □

Задачи

1. Сравнете Теореми 8.2 и 8.3 съответно с теоремите за гладки и за изпъкнали задачи.

2. Изведете с помощта на Теорема 8.2 необходимо условие за локален минимум за екстремалната задача

$$\left| \begin{array}{l} f_0(x) \longrightarrow \inf \\ F(x) = 0, \\ f_i(x) \leq 0, \quad i = 1, 2, \ldots, n, \end{array} \right.$$

   при съответните предположения върху изображенията $f_i : X \to \mathbb{R}$, $i = 0, 1, \ldots, n$, и $F : X \to Y$, където X и Y са банахови пространства.

9 Принцип за максимума — формулировка и основни частни случаи

Сега ще разгледаме един клас задачи от оптималното управление. Нека са дадени изображенията

$$\begin{aligned} &f : V \to \mathbb{R}, \\ &\varphi : V \to \mathbb{R}^n, \\ &b_i : V_i \to \mathbb{R}^{s_i}, \quad i = 0, 1. \end{aligned}$$

където $V \subseteq \mathbb{R}^{m+n+1}$ и $V_i \subseteq \mathbb{R}^{n+1}$, $i = 0, 1$. Нека също $U \subseteq \mathbb{R}^m$.

Разглеждаме екстремалната задача

$$\left| \begin{aligned} &\mathcal{F}(x, u) := \int_{t_0}^{t_1} f(t, x(t), u(t))\, dt \longrightarrow \inf \\ &\dot{x}(t) = \varphi(t, x(t), u(t)), \quad t \in [t_0, t_1], \qquad (*) \\ &u(t) \in U, \quad t \in [t_0, t_1], \qquad (**) \\ &b_0(t_0, x(t_0)) = 0, \quad b_1(t_1, x(t_1)) = 0. \qquad (***) \end{aligned} \right. \tag{9.1}$$

Тук $x(t) = (x_1(t), x_2(t), \ldots, x_n(t))$ е векторно-значна функция, чиито координатни функции предполагаме абсолютно непрекъснати и удовлетворяващи $(*)$ почти навсякъде относно лебеговата мярка. Тя се нарича *фазова траектория*.

Векторно-значната функция $u(t) = (u_1(t), u_2(t), \ldots, u_m(t))$ е с координатни функции, които са ограничени, измерими (отново по Лебег) и удовлетворяват $(**)$ почти навсякъде. Тя се нарича *допустимо управление*. Тук идеята е, че имаме контрол върху моделирания процес посредством u. Желаем така да определим u, че минимумът на функционала $\mathcal{F}(x, u)$ по траекториите x да е възможно най-малък.

Интервалът $[t_0, t_1]$ в общия случай не е фиксиран. Така той също се явява променлива в екстремалната задача.

Дефиниция 9.1. Всяка наредена тройка $(x(t), u(t), [t_0, t_1])$, чиито компоненти удовлетворяват условията $(*)$-$(***)$, наричаме *допустим процес* за задачата (9.1).

Дефиниция 9.2. Казваме, че допустимият процес $(\widetilde{x}(t), \widetilde{u}(t), [\widetilde{t}_0, \widetilde{t}_1])$ е оптимален за задачата (9.1), ако съществува $\delta > 0$ такова, че

$$\mathcal{F}(\widetilde{x}, \widetilde{u}) \leq \mathcal{F}(x, u) \tag{9.2}$$

за всеки друг допустим процес $(x(t), u(t), [t_0, t_1])$ с

$$|t_0 - \widetilde{t}_0| < \delta, \quad |t_1 - \widetilde{t}_1| < \delta, \quad |x(t) - \widetilde{x}(t)| < \delta,\ t \in [t_0, t_1] \cap [\widetilde{t}_0, \widetilde{t}_1].$$

Казваме още, че допустимият процес $(\widetilde{x}(t), \widetilde{u}(t), [\widetilde{t}_0, \widetilde{t}_1])$ реализира силен локален минимум за задачата (9.1).

Да споменем изрично, че (9.2) има вида

$$\int_{\widetilde{t}_0}^{\widetilde{t}_1} f(t, \widetilde{x}(t), \widetilde{u}(t))\, dt \leq \int_{t_0}^{t_1} f(t, x(t), u(t))\, dt.$$

Ще формулираме необходими условия, при които даден процес е оптимален за задачата (9.1). Те се наричат Принцип за максимума или още Принцип на Понтрягин. Съществуват две еквивалентни формулировки. Едната се нарича форма на Хамилтън, а другата — форма на Лагранж. Доказателството ще бъде изложено в следващата лекция.

9.1 Форма на Хамилтън

За $(t, x, u) \in V$, $p \in \mathbb{R}^n$ и $\lambda \geq 0$ дефинираме функцията

$$H(t, x, u, p, \lambda) := p \cdot \varphi(t, x, u) - \lambda f(t, x, u).$$

Тя се нарича *функция на Понтрягин* за задачата (9.1). А функцията

$$\mathcal{H}(t, x, p, \lambda) := \sup_{u \in U} H(t, x, u, p, \lambda)$$

се нарича *хамилтониан*.

Теорема 9.3 (Принцип за максимума във форма на Хамилтън)**.** *Нека $f, \varphi \in C(V)$, $b_i \in C^1(V_i)$, $i = 0, 1$, където $V \subseteq \mathbb{R}^{n+m+1}$ и $V_i \subseteq \mathbb{R}^{n+1}$ са отворени. Нека f и φ са непрекъснато диференцируеми по $t, x_1, x_2, \ldots, x_n$. Накрая нека φ е ограничена върху ограничените подмножества на V. Ако $(\widetilde{x}(t), \widetilde{u}(t), [\widetilde{t}_0, \widetilde{t}_1])$ е оптимален процес за задачата (9.1), то съществуват $\lambda \geq 0$, $\ell_0 \in \mathbb{R}^{s_0}$, $\ell_1 \in \mathbb{R}^{s_1}$ и абсолютно непрекъсната векторнозначна функция $p : [\widetilde{t}_0, \widetilde{t}_1] \to \mathbb{R}^n$, не всичките нула, такива, че*

(i) *имаме*

$$\begin{aligned} \dot{p}(t) &= -H'_x(t, \widetilde{x}(t), \widetilde{u}(t), p(t), \lambda) \quad \text{п.н. в} \quad [\widetilde{t}_0, \widetilde{t}_1], \\ p(\widetilde{t}_i) &= (-1)^i b'_{i_x}(\widetilde{t}_i, \widetilde{x}(\widetilde{t}_i))^* \ell_i, \quad i = 0, 1; \end{aligned}$$

(ii) *п.н. в $[\widetilde{t}_0, \widetilde{t}_1]$ имаме*

$$\begin{aligned} H(t, \widetilde{x}(t), \widetilde{u}(t), p(t), \lambda) &= \max_{u \in U} H(t, \widetilde{x}(t), u, p(t), \lambda) \\ &= \mathcal{H}(t, \widetilde{x}(t), p(t), \lambda); \end{aligned}$$

(iii) *функцията $\mathcal{H}(t, \widetilde{x}(t), p(t), \lambda)$ е непрекъсната в $[\widetilde{t}_0, \widetilde{t}_1]$ и удовлетворява съотношенията*

$$\mathcal{H}(\widetilde{t}_i, \widetilde{x}(\widetilde{t}_i), p(\widetilde{t}_i), \lambda) = (-1)^{i+1} b'_{i_t}(\widetilde{t}_i, \widetilde{x}(\widetilde{t}_i)) \cdot \ell_i \quad i = 0, 1.$$

Все така λ, ℓ_0, ℓ_1 и p се наричат *множители на Лагранж*. Диференциалното уравнение в (i) се нарича *спрегнато уравнение*, а граничните условия — *условия за трансверзалност*.

Нека отбележим, че

$$H'_x(t, x, u, p, \lambda) = \varphi'_x(t, x, u)^* p - \lambda f'_x(t, x, u)$$

и

$$H'_t(t, x, u, p, \lambda) = p \cdot \varphi'_t(t, x, u) - \lambda f'_t(t, x, u).$$

В сила е още следната формула

$$\mathcal{H}(t, \widetilde{x}(t), p(t), \lambda) = b'_{1_t}(\widetilde{t}_1, \widetilde{x}(\widetilde{t}_1)) \cdot \ell_1 - \int_t^{\widetilde{t}_1} H'_t(\tau, \widetilde{x}(\tau), \widetilde{u}(\tau), p(\tau), \lambda)\, d\tau. \tag{9.3}$$

9.2 Форма на Лагранж

За $t \in [t_0, t_1]$, $x \in AC([t_0, t_1], \mathbb{R}^n)$ и $u \in U$ такива, че $(t, x, u) \in V$ п.н., и $p \in \mathbb{R}^n$ и $\lambda \in \mathbb{R}$ дефинираме функцията

$$L(t, x, \dot{x}, u, \lambda, p) := p \cdot (\dot{x} - \varphi(t, x, u)) + \lambda f(t, x, u).$$

Тя се нарича *лагранжиан* на задачата (9.1). А функцията

$$\mathcal{L}(t_0, t_1, x, u, \lambda, p, \ell_0, \ell_1) := \int_{t_0}^{t_1} L(t, x(t), \dot{x}(t), u(t), \lambda, p(t))\, dt \\ + \ell_0 \cdot b_0(t_0, x(t_0)) + \ell_1 \cdot b_1(t_1, x(t_1))$$

се нарича *функция на Лагранж*.

Теорема 9.4 (Принцип за максимума във форма на Лагранж)**.** *Нека $f, \varphi \in C(V)$, $b_i \in C^1(V_i)$, $i = 0, 1$, където $V \subseteq \mathbb{R}^{n+m+1}$ и $V_i \subseteq \mathbb{R}^{n+1}$ са отворени. Нека f и φ са непрекъснато диференцируеми по $t, x_1, x_2, \dots, x_n$. Накрая нека φ е ограничена върху ограничените подмножества на V. Ако $(\widetilde{x}(t), \widetilde{u}(t), [\widetilde{t}_0, \widetilde{t}_1])$ е оптимален процес за задачата (9.1), то съществуват $\lambda \geq 0$, $\ell_0 \in \mathbb{R}^{s_0}$, $\ell_1 \in \mathbb{R}^{s_1}$, и абсолютно непрекъсната векторнозначна функция $p : [\widetilde{t}_0, \widetilde{t}_1] \to \mathbb{R}^n$, не всичките нула, такива, че*

(i) *имаме*

$$\frac{d}{dt} L'_{\dot{x}}(t, \widetilde{x}(t), \dot{\widetilde{x}}(t), \widetilde{u}(t), \lambda, p(t)) = L'_x(t, \widetilde{x}(t), \dot{\widetilde{x}}(t), \widetilde{u}(t), \lambda, p(t)) \\ \text{п.н. в} \quad [\widetilde{t}_0, \widetilde{t}_1],$$

$$L'_{\dot{x}}(\widetilde{t}_i, \widetilde{x}(\widetilde{t}_i), \dot{\widetilde{x}}(\widetilde{t}_i), \widetilde{u}(\widetilde{t}_i), \lambda, p(\widetilde{t}_i)) = (-1)^i b'_{i_x}(\widetilde{t}_i, \widetilde{x}(\widetilde{t}_i))^* \ell_i, \quad i = 0, 1;$$

(ii) *п.н. в* $[\widetilde{t}_0, \widetilde{t}_1]$ *имаме*

$$L(t, \widetilde{x}(t), \dot{\widetilde{x}}(t), \widetilde{u}(t), \lambda, p(t)) = \min_{u \in U} L(t, \widetilde{x}(t), \dot{\widetilde{x}}(t), u, \lambda, p(t));$$

(iii) *функцията* $\mathcal{L}(t_0, t_1) := \mathcal{L}(t_0, t_1, \widetilde{x}, \widetilde{u}, \lambda, p, \ell_0, \ell_1)$, $t_0, t_1 \in [\widetilde{t}_0, \widetilde{t}_1]$, *е диференцируема по* t_0 *отдясно в точката* $\widetilde{t}_0$ *и по* t_1 *отляво в точката* $\widetilde{t}_1$, *при това*

$$\mathcal{L}'_{t_0}(\widetilde{t}_0 + 0, t_1) = 0, \quad \mathcal{L}'_{t_1}(t_0, \widetilde{t}_1 - 0) = 0.$$

9.3 Еквивалентност

Ще установим еквивалентността на двете формулировки на Принципа на Понтрягин за максимума. За краткост няма да пишем аргументите на функцията на Понтрягин и на лагранжиана, освен когато може да възникне неяснота.

Имаме

$$L = p \cdot \dot{x} - H. \tag{9.4}$$

Следователно

$$L'_{\dot{x}} = p, \quad L'_x = -H'_x.$$

Оттук непосредствено следва, че условията (i) и (ii) в двете теореми са съответно еквивалентни помежду си.

Остава да установим еквивалентността на третите свойства. В доказателството на Теорема 9.3 ще покажем, че непрекъснатостта на $\mathcal{H}$ следва от (i) и (ii) (в същата теорема), а тогава и от Теорема 9.4(i)-(ii). Остава да покажем съоветно еквивалентността на съотношенията относно точките $\widetilde{t}_0$ и $\widetilde{t}_1$.

Теорема 9.4(ii) и (9.4) влекат

$$\begin{aligned}\mathcal{L}(t_0, t_1, \widetilde{x}, \widetilde{u}, \lambda, p, \ell_0, \ell_1) = \int_{t_0}^{t_1} \big[p(t) \cdot \dot{\widetilde{x}}(t) - \mathcal{H}(t, \widetilde{x}(t), p(t), \lambda)\big]\, dt \\ + \ell_0 \cdot b_0(t_0, \widetilde{x}(t_0)) + \ell_1 \cdot b_1(t_1, \widetilde{x}(t_1)).\end{aligned}$$

Следователно за $\varepsilon > 0$ (достатъчно малко) имаме

$$\begin{aligned}\mathcal{L}(\widetilde{t}_0 + \varepsilon, t_1) - \mathcal{L}(\widetilde{t}_0, t_1) = \ell_0 \cdot \big[b_0(\widetilde{t}_0 + \varepsilon, \widetilde{x}(\widetilde{t}_0 + \varepsilon)) - b_0(\widetilde{t}_0, \widetilde{x}(\widetilde{t}_0))\big] \\ - \int_{\widetilde{t}_0}^{\widetilde{t}_0 + \varepsilon} \big[p(t) \cdot \dot{\widetilde{x}}(t) - \mathcal{H}(t, \widetilde{x}(t), p(t), \lambda)\big]\, dt.\end{aligned} \tag{9.5}$$

Като използваме, че изображението b_0 е диференцируемо по променливата (t, x) (вж. 2.3(б)), извеждаме представянето

$$\begin{aligned}b_0(\widetilde{t}_0 + \varepsilon, \widetilde{x}(\widetilde{t}_0 + \varepsilon)) - b_0(\widetilde{t}_0, \widetilde{x}(\widetilde{t}_0)) = \varepsilon b'_{0_t}(\widetilde{t}_0, \widetilde{x}(\widetilde{t}_0)) \\ + b'_{0_x}(\widetilde{t}_0, \widetilde{x}(\widetilde{t}_0))\big[\widetilde{x}(\widetilde{t}_0 + \varepsilon) - \widetilde{x}(\widetilde{t}_0)\big] + o(\max\{\varepsilon, |\widetilde{x}(\widetilde{t}_0 + \varepsilon) - \widetilde{x}(\widetilde{t}_0)|\}).\end{aligned} \tag{9.6}$$

Тъй като $\widetilde{x}$ е непрекъсната и $\widetilde{u}$ е ограничена, то множеството от точки в $\mathbb{R}^{n+m+1}$ $\{(t, \widetilde{x}(t), \widetilde{u}(t)) : t \in [\widetilde{t}_0, \widetilde{t}_1]\}$ е ограничено. Като, вземем предвид, че φ е ограничена върху ограничените подмножества на V, извеждаме от $(*)$, че $\dot{\widetilde{x}} \in L_\infty[\widetilde{t}_0, \widetilde{t}_1]$. Оттук получаваме, че

$$
\begin{aligned}
|\widetilde{x}(\widetilde{t}_0 + \varepsilon) - \widetilde{x}(\widetilde{t}_0)| &= \left| \int_{\widetilde{t}_0}^{\widetilde{t}_0+\varepsilon} \dot{\widetilde{x}}(t)\, dt \right| \\
&\leq \int_{\widetilde{t}_0}^{\widetilde{t}_0+\varepsilon} |\dot{\widetilde{x}}(t)|\, dt \leq \varepsilon \|\dot{\widetilde{x}}\|_{L_\infty}.
\end{aligned} \tag{9.7}
$$

Предвид това, (9.6) придобива вида

$$
\begin{aligned}
b_0(\widetilde{t}_0 + \varepsilon, \widetilde{x}(\widetilde{t}_0 + \varepsilon)) - b_0(\widetilde{t}_0, \widetilde{x}(\widetilde{t}_0)) &= \varepsilon b'_{0_t}(\widetilde{t}_0, \widetilde{x}(\widetilde{t}_0)) \\
&+ b'_{0_x}(\widetilde{t}_0, \widetilde{x}(\widetilde{t}_0))\big[\widetilde{x}(\widetilde{t}_0 + \varepsilon) - \widetilde{x}(\widetilde{t}_0)\big] + o(\varepsilon),
\end{aligned}
$$

откъдето получаваме

$$
\begin{aligned}
\ell_0 \cdot \big[b_0(\widetilde{t}_0 + \varepsilon, \widetilde{x}(\widetilde{t}_0 + \varepsilon)) - b_0(\widetilde{t}_0, \widetilde{x}(\widetilde{t}_0))\big] &= \varepsilon \ell_0 \cdot b'_{0_t}(\widetilde{t}_0, \widetilde{x}(\widetilde{t}_0)) \\
&+ b'_{0_x}(\widetilde{t}_0, \widetilde{x}(\widetilde{t}_0))^* \ell_0 \cdot \big[\widetilde{x}(\widetilde{t}_0 + \varepsilon) - \widetilde{x}(\widetilde{t}_0)\big] + o(\varepsilon).
\end{aligned} \tag{9.8}
$$

По-нататък очевидно

$$
\begin{aligned}
\int_{\widetilde{t}_0}^{\widetilde{t}_0+\varepsilon} p(t) \cdot \dot{\widetilde{x}}(t)\, dt &= \int_{\widetilde{t}_0}^{\widetilde{t}_0+\varepsilon} [p(t) - p(\widetilde{t}_0)] \cdot \dot{\widetilde{x}}(t)\, dt \\
&\qquad + p(\widetilde{t}_0) \cdot \int_{\widetilde{t}_0}^{\widetilde{t}_0+\varepsilon} \dot{\widetilde{x}}(t)\, dt \\
&= \int_{\widetilde{t}_0}^{\widetilde{t}_0+\varepsilon} [p(t) - p(\widetilde{t}_0)] \cdot \dot{\widetilde{x}}(t)\, dt + p(\widetilde{t}_0) \cdot [\widetilde{x}(\widetilde{t}_0 + \varepsilon) - \widetilde{x}(\widetilde{t}_0)].
\end{aligned} \tag{9.9}
$$

Оценяваме първото събираемо горе, като използваме неравенството на Коши, непрекъснатостта на $p(t)$ и (9.7):

$$
\begin{aligned}
\left| \int_{\widetilde{t}_0}^{\widetilde{t}_0+\varepsilon} [p(t) - p(\widetilde{t}_0)] \cdot \dot{\widetilde{x}}(t)\, dt \right| &\leq \int_{\widetilde{t}_0}^{\widetilde{t}_0+\varepsilon} |[p(t) - p(\widetilde{t}_0)] \cdot \dot{\widetilde{x}}(t)|\, dt \\
&\leq \int_{\widetilde{t}_0}^{\widetilde{t}_0+\varepsilon} |p(t) - p(\widetilde{t}_0)|\, |\dot{\widetilde{x}}(t)|\, dt \\
&\leq \max_{t \in [\widetilde{t}_0, \widetilde{t}_0+\varepsilon]} |p(t) - p(\widetilde{t}_0)| \int_{\widetilde{t}_0}^{\widetilde{t}_0+\varepsilon} |\dot{\widetilde{x}}(t)|\, dt \\
&= o(\varepsilon).
\end{aligned} \tag{9.10}
$$

Сега от (9.9) и (9.10) следва

$$\int_{\widetilde{t}_0}^{\widetilde{t}_0+\varepsilon} p(t)\cdot\dot{\widetilde{x}}(t)\,dt = p(\widetilde{t}_0)\cdot[\widetilde{x}(\widetilde{t}_0+\varepsilon)-\widetilde{x}(\widetilde{t}_0)] + o(\varepsilon). \tag{9.11}$$

Накрая, като вземем предвид, че $\mathcal{H}(t,\widetilde{x}(t),p(t),\lambda)$ е непрекъсната по t, получаваме

$$\begin{aligned}\int_{\widetilde{t}_0}^{\widetilde{t}_0+\varepsilon} \mathcal{H}(t,\widetilde{x}(t),p(t),\lambda)\,dt &= \varepsilon\,\mathcal{H}(\widetilde{t}_0,\widetilde{x}(\widetilde{t}_0),p(\widetilde{t}_0),\lambda)\\ &\quad+\int_{\widetilde{t}_0}^{\widetilde{t}_0+\varepsilon}\big[\mathcal{H}(t,\widetilde{x}(t),p(t),\lambda)-\mathcal{H}(\widetilde{t}_0,\widetilde{x}(\widetilde{t}_0),p(\widetilde{t}_0),\lambda)\big]\,dt\\ &= \varepsilon\,\mathcal{H}(\widetilde{t}_0,\widetilde{x}(\widetilde{t}_0),p(\widetilde{t}_0),\lambda)+o(\varepsilon).\end{aligned} \tag{9.12}$$

От (9.5), (9.8), (9.11), (9.12) и граничното условие в точката $\widetilde{t}_0$ в Теорема 9.3(i) получаваме

$$\mathcal{L}(\widetilde{t}_0+\varepsilon,t_1)-\mathcal{L}(\widetilde{t}_0,t_1)=\varepsilon\big[\ell_0\cdot b'_{0_t}(\widetilde{t}_0,\widetilde{x}(\widetilde{t}_0))+\mathcal{H}(\widetilde{t}_0,\widetilde{x}(\widetilde{t}_0),p(\widetilde{t}_0),\lambda)\big]+o(\varepsilon).$$

Оттук непосредствено следва еквивалентността на съотношенията в (iii) в двете теореми относно $\widetilde{t}_0$. Аналогично се разглежда случая в $\widetilde{t}_1$.

9.4 Коментари и частни случаи

Принципът за максимума е реализация на Принципа на Лагранж, който, да припомним, гласи, че необходими условия за условен екстремум съвпадат с необходими условия за екстремум на функцията на Лагранж при условията, които не са отразени в нея.

Нека сега разгледаме няколко съществени частни случая на Принципа за максимума. Ще направим това във формата на Хамилтън.

(a) Фиксирана стойност на траекторията в край: $x(t_0)=x_0\in\mathbb{R}^n$, като t_0 остава свободен. Тогава $b_0(t,x):=x-x_0$ и следователно $b'_{0_x}(t,x)$ е идентитетът в $\mathbb{R}^n$ (квадратната матрица $n\times n$ с единици по главния диагонал, а останалите елементи — нули) и $b'_{0_t}(t,x)\equiv\mathbf{0}:=(0,0,\dots,0)\in\mathbb{R}^n$. Граничното условие в (i) за $i=0$ придобива вида $p(\widetilde{t}_0)=\ell_0$, а съответното условие в (iii) — $\mathcal{H}(\widetilde{t}_0,\widetilde{x}(\widetilde{t}_0),p(\widetilde{t}_0),\lambda)=0$. При това не е възможно $\ell_0\neq\mathbf{0}$, а всички останали множители на Лагранж да са нула, защото $p(t)\equiv\mathbf{0}$ влече $\ell_0=\mathbf{0}$. Следователно множителят на Лагранж ℓ_0 не носи информация и можем да не го въвеждаме. Така Теорема 9.3 влече следното необходимо условие за решението на екстремалната задача

$$
(9.13)\qquad \left|\begin{array}{ll}
\mathcal{F}(x,u) := \displaystyle\int_{t_0}^{t_1} f(t,x(t),u(t))\,dt \longrightarrow \inf & \\
\dot{x}(t) = \varphi(t,x(t),u(t)), \quad t \in [t_0,t_1], & (*) \\
u(t) \in U, \quad t \in [t_0,t_1], & (**) \\
x(t_0) = x_0, \quad b_1(t_1,x(t_1)) = 0, & (***)
\end{array}\right.
$$

където $x_0 \in \mathbb{R}^n$ е фиксирано.

Теорема 9.5 (Принцип за максимума във форма на Хамилтън)**.** *Нека $f, \varphi \in C(V)$, $b_1 \in C^1(V_1)$, където $V \subseteq \mathbb{R}^{n+m+1}$ и $V_1 \subseteq \mathbb{R}^{n+1}$ са отворени. Нека f и φ са непрекъснато диференцируеми по $t, x_1, x_2, \ldots, x_n$. Накрая нека φ е ограничена върху ограничените подмножества на V. Ако $(\widetilde{x}(t), \widetilde{u}(t), [\widetilde{t}_0, \widetilde{t}_1])$ е оптимален процес за задачата (9.13), то съществуват $\lambda \geq 0$, $\ell_1 \in \mathbb{R}^{s_1}$ и абсолютно непрекъсната векторно-значна функция $p : [\widetilde{t}_0, \widetilde{t}_1] \to \mathbb{R}^n$, не всичките нула, такива, че*

(i) *имаме*

$$
\begin{aligned}
&\dot{p}(t) = -H'_x(t, \widetilde{x}(t), \widetilde{u}(t), p(t), \lambda) \quad \text{п.н. в} \quad [\widetilde{t}_0, \widetilde{t}_1],\\
&p(\widetilde{t}_1) = -b'_{1_x}(\widetilde{t}_1, \widetilde{x}(\widetilde{t}_1))^* \ell_1;
\end{aligned}
$$

(ii) *п.н. в $[\widetilde{t}_0, \widetilde{t}_1]$ имаме*

$$
\begin{aligned}
H(t, \widetilde{x}(t), \widetilde{u}(t), p(t), \lambda) &= \max_{u \in U} H(t, \widetilde{x}(t), u, p(t), \lambda)\\
&= \mathcal{H}(t, \widetilde{x}(t), p(t), \lambda);
\end{aligned}
$$

(iii) *функцията $\mathcal{H}(t, \widetilde{x}(t), p(t), \lambda)$ е непрекъсната в $[\widetilde{t}_0, \widetilde{t}_1]$ и удовлетворява съотношенията:*

$$
\begin{aligned}
&\mathcal{H}(\widetilde{t}_0, \widetilde{x}(\widetilde{t}_0), p(\widetilde{t}_0), \lambda) = 0,\\
&\mathcal{H}(\widetilde{t}_1, \widetilde{x}(\widetilde{t}_1), p(\widetilde{t}_1), \lambda) = b'_{1_t}(\widetilde{t}_1, \widetilde{x}(\widetilde{t}_1)) \cdot \ell_1.
\end{aligned}
$$

(б) Фиксиран времеви момент: t_0 е фиксирано, като стойността на $x(t)$ в t_0 остава свободна. Тогава $b_0(t,x) := t - t_0$ и следователно $b'_{0_x}(t,x) \equiv \mathbf{0}$ и $b'_{0_t}(t,x) \equiv 1$. Също така сега имаме $\widetilde{t}_0 = t_0$. Условията в (i) и (iii) в t_0 придобиват вида $p(t_0) = \mathbf{0}$ и $\mathcal{H}(t_0, \widetilde{x}(t_0), p(t_0), \lambda) = -\ell_0$, като последното е тривиално. Отново ℓ_0 не носи информация и можем да не го въвеждаме. В този случай екстремалната задача (9.1) има вида

$$
(9.14)\qquad \left|\begin{array}{ll}
\mathcal{F}(x,u) := \displaystyle\int_{t_0}^{t_1} f(t,x(t),u(t))\,dt \longrightarrow \inf & \\
\dot{x}(t) = \varphi(t,x(t),u(t)), \quad t \in [t_0,t_1], & (*) \\
u(t) \in U, \quad t \in [t_0,t_1], & (**) \\
b_1(t_1,x(t_1)) = 0, & (***)
\end{array}\right.
$$

като t_0 е фиксирано (стойността на $x(t)$ в t_0 е свободна). От Теорема 9.3 следва необходимото условие:

Теорема 9.6 (Принцип за максимума във форма на Хамилтън)**.** *Нека $f, \varphi \in C(V)$, $b_1 \in C^1(V_1)$, където $V \subseteq \mathbb{R}^{n+m+1}$ и $V_1 \subseteq \mathbb{R}^{n+1}$ са отворени. Нека f и φ са непрекъснато диференцируеми по $t, x_1, x_2, \dots, x_n$. Накрая нека φ е ограничена върху ограничените подмножества на V. Ако $(\widetilde{x}(t), \widetilde{u}(t), [t_0, \widetilde{t}_1])$ е оптимален процес за задачата* (9.14)*, то съществуват $\lambda \geq 0$, $\ell_1 \in \mathbb{R}^{s_1}$ и абсолютно непрекъсната векторно-значна функция $p : [t_0, \widetilde{t}_1] \to \mathbb{R}^n$, не всичките нула, такива, че*

(i) *имаме*

$$\begin{aligned} &\dot{p}(t) = -H'_x(t, \widetilde{x}(t), \widetilde{u}(t), p(t), \lambda) \quad \text{п.н. в} \quad [t_0, \widetilde{t}_1], \\ &p(t_0) = \mathbf{0}, \\ &p(\widetilde{t}_1) = -b'_{1_x}(\widetilde{t}_1, \widetilde{x}(\widetilde{t}_1))^* \ell_1; \end{aligned}$$

(ii) *п.н. в $[t_0, \widetilde{t}_1]$ имаме*

$$\begin{aligned} H(t, \widetilde{x}(t), \widetilde{u}(t), p(t), \lambda) &= \max_{u \in U} H(t, \widetilde{x}(t), u, p(t), \lambda) \\ &= \mathcal{H}(t, \widetilde{x}(t), p(t), \lambda); \end{aligned}$$

(iii) *функцията $\mathcal{H}(t, \widetilde{x}(t), p(t), \lambda)$ е непрекъсната в $[t_0, \widetilde{t}_1]$ и удовлетворява съотношението*

$$\mathcal{H}(\widetilde{t}_1, \widetilde{x}(\widetilde{t}_1), p(\widetilde{t}_1), \lambda) = b'_{1_t}(\widetilde{t}_1, \widetilde{x}(\widetilde{t}_1)) \cdot \ell_1.$$

(в) Фиксиран край: t_0 е фиксирано и $x(t_0) = x_0 \in \mathbb{R}^n$ също е фиксирано. Тогава $b_0(t, x) := (t - t_0, x - x_0)$. Този случай представлява своеобразна комбинация от първите два. Отново получаваме, че множителят на Лагранж ℓ_0 е излишен. По-подробно, имаме

$$b'_{0_x}(t, x) = \underbrace{\begin{pmatrix} 0 & 0 & \dots & 0 \\ 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 \end{pmatrix}}_{n \text{ стълба}} \left.\vphantom{\begin{pmatrix} 0 \\ 1 \\ 0 \\ \dots \\ 0 \end{pmatrix}}\right\} \quad n+1 \text{ реда}$$

и $b'_{0_t}(t, x) = (1, 0, \dots, 0) \in \mathbb{R}^{n+1}$. Нека $\ell_0 := (\ell_{0,0}, \ell_{0,1}, \dots, \ell_{0,n})$, където $\ell_{0,j} \in \mathbb{R}$, $j = 0, 1, \dots, n$. Тогава, взимайки предвид, че $b'_{0_x}(t_0, x_0)^*$ се получава от $b'_{0_x}(t_0, x_0)$ чрез транспониране, получаваме

$$b'_{0_x}(t_0, x_0)^* \ell_0 = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix} \begin{pmatrix} \ell_{0,0} \\ \ell_{0,1} \\ \vdots \\ \ell_{0,n} \end{pmatrix} = \begin{pmatrix} \ell_{0,1} \\ \ell_{0,2} \\ \vdots \\ \ell_{0,n} \end{pmatrix}.$$

Следователно $p(t_0) = (\ell_{0,1}, \ell_{0,2}, \dots, \ell_{0,n})$.

Имаме също $b'_{0_t}(t,x) \cdot \ell_0 = \ell_{0,0}$ и условието в (iii) при $i = 0$ има вида $\mathcal{H}(t_0, x_0, p(t_0), \lambda) = -\ell_{0,0}$.

Ако допуснем, че $p(t) \equiv \mathbf{0}$, получаваме, че $\ell_{0,j} = 0$ за $j = 1, 2, \dots, n$. Ако допуснем още, че и $\lambda = 0$, то $H(t, x, u, 0, 0) \equiv 0$ и следователно $\mathcal{H}(t, x, 0, 0) \equiv 0$ и тогава $\ell_{0,0} = 0$. Така, ако допуснем, че множителите на Лагранж, които съответстват на дадено решение на екстремалната задача (9.1), λ, ℓ_1 и p са нула, ще получим и че $\ell_0 = \mathbf{0} \in \mathbb{R}^{n+1}$, което е противоречие.

Установихме, че не е възможно $\ell_0 \neq \mathbf{0}$, а останалите множители на Лагранж да са нула. Следователно множителят на Лагранж ℓ_0 не носи информация и можем да не го въвеждаме.

Да формулираме Принципа на максимума в разглеждания случай. Задачата (9.1) сега има вида

$$
\text{(9.15)} \qquad \left|\begin{array}{ll}
\mathcal{F}(x,u) := \displaystyle\int_{t_0}^{t_1} f(t, x(t), u(t))\, dt \longrightarrow \inf & \\
\dot{x}(t) = \varphi(t, x(t), u(t)), \quad t \in [t_0, t_1], & (*) \\
u(t) \in U, \quad t \in [t_0, t_1], & (**) \\
x(t_0) = x_0, \quad b_1(t_1, x(t_1)) = 0, & (***)
\end{array}\right.
$$

като $t_0 \in \mathbb{R}$ и $x_0 \in \mathbb{R}^n$ са фиксирани. От Теорема 9.3 следва необходимото условие:

Теорема 9.7 (Принцип за максимума във форма на Хамилтън)**.** *Нека $f, \varphi \in C(V)$, $b_1 \in C^1(V_1)$, където $V \subseteq \mathbb{R}^{n+m+1}$ и $V_1 \subseteq \mathbb{R}^{n+1}$ са отворени. Нека f и φ са непрекъснато диференцируеми по $t, x_1, x_2, \dots, x_n$. Накрая нека φ е ограничена върху ограничените подмножества на V. Ако $(\widetilde{x}(t), \widetilde{u}(t), [t_0, \widetilde{t}_1])$ е оптимален процес за задачата (9.14), то съществуват $\lambda \geq 0$, $\ell_1 \in \mathbb{R}^{s_1}$ и абсолютно непрекъсната векторно-значна функция $p : [t_0, \widetilde{t}_1] \to \mathbb{R}^n$, не всичките нула, такива, че*

(i) *имаме*

$$
\begin{aligned}
&\dot{p}(t) = -H'_x(t, \widetilde{x}(t), \widetilde{u}(t), p(t), \lambda) \quad \text{п.н. в} \quad [t_0, \widetilde{t}_1], \\
&p(\widetilde{t}_1) = -b'_{1_x}(\widetilde{t}_1, \widetilde{x}(\widetilde{t}_1))^* \ell_1;
\end{aligned}
$$

(ii) *п.н. в $[t_0, \widetilde{t}_1]$ имаме*

$$
\begin{aligned}
H(t, \widetilde{x}(t), \widetilde{u}(t), p(t), \lambda) &= \max_{u \in U} H(t, \widetilde{x}(t), u, p(t), \lambda) \\
&= \mathcal{H}(t, \widetilde{x}(t), p(t), \lambda);
\end{aligned}
$$

(iii) *функцията $\mathcal{H}(t, \widetilde{x}(t), p(t), \lambda)$ е непрекъсната в $[t_0, \widetilde{t}_1]$ и удовлетворява съотношението*

$$
\mathcal{H}(\widetilde{t}_1, \widetilde{x}(\widetilde{t}_1), p(\widetilde{t}_1), \lambda) = b'_{1_t}(\widetilde{t}_1, \widetilde{x}(\widetilde{t}_1)) \cdot \ell_1.
$$

(г) Свободен край: не са поставени ограничения нито върху t_0, нито върху $x(t_0)$, т.е. разглеждаме задачата

$$
(9.16)\qquad \left|\begin{array}{ll}
\mathcal{F}(x,u) := \displaystyle\int_{t_0}^{t_1} f(t,x(t),u(t))\,dt \longrightarrow \inf & \\
\dot{x}(t) = \varphi(t,x(t),u(t)), \quad t \in [t_0,t_1], & (*) \\
u(t) \in U, \quad t \in [t_0,t_1], & (**) \\
b_1(t_1,x(t_1)) = 0. & (***)
\end{array}\right.
$$

В този случай Теорема 9.3 непосредствено влече необходими условия за решение, като положим $b_0(t,x) := 0$. Тогава получаваме, че решенията на задача (9.16) трябва да удовлетворяват условията

$$
p(\widetilde{t}_0) = \mathbf{0} \quad \text{и} \quad \mathcal{H}(\widetilde{t}_0, \widetilde{x}(\widetilde{t}_0), p(\widetilde{t}_0), \lambda) = 0.
$$

Множителят на Лагранж ℓ_0 не се среща в нито едно необходимо условие, но и не можем да изключим случая, в който всички останали множители на Лагранж са нула. С малка модификация на доказателството на Теорема 9.3 се установява следната естествена форма на Принципа за максимума за задачата (9.16).

Теорема 9.8 (Принцип за максимума във форма на Хамилтън). *Нека $f, \varphi \in C(V)$, $b_1 \in C^1(V_1)$, където $V \subseteq \mathbb{R}^{n+m+1}$ и $V_1 \subseteq \mathbb{R}^{n+1}$ са отворени. Нека f и φ са непрекъснато диференцируеми по $t, x_1, x_2, \ldots, x_n$. Накрая нека φ е ограничена върху ограничените подмножества на V. Ако $(\widetilde{x}(t), \widetilde{u}(t), [\widetilde{t}_0, \widetilde{t}_1])$ е оптимален процес за задачата* (9.16)*, то съществуват $\lambda \geq 0$, $\ell_1 \in \mathbb{R}^{s_1}$ и абсолютно непрекъсната векторно-значна функция $p : [\widetilde{t}_0, \widetilde{t}_1] \to \mathbb{R}^n$, не всичките нула, такива, че*

(i) *имаме*

$$
\begin{aligned}
&\dot{p}(t) = -H'_x(t, \widetilde{x}(t), \widetilde{u}(t), p(t), \lambda) \quad \text{п.н. в} \quad [\widetilde{t}_0, \widetilde{t}_1],\\
&p(\widetilde{t}_0) = \mathbf{0},\\
&p(\widetilde{t}_1) = -b'_{1_x}(\widetilde{t}_1, \widetilde{x}(\widetilde{t}_1))^* \ell_1;
\end{aligned}
$$

(ii) *п.н. в $[\widetilde{t}_0, \widetilde{t}_1]$ имаме*

$$
\begin{aligned}
H(t, \widetilde{x}(t), \widetilde{u}(t), p(t), \lambda) &= \max_{u \in U} H(t, \widetilde{x}(t), u, p(t), \lambda)\\
&= \mathcal{H}(t, \widetilde{x}(t), p(t), \lambda);
\end{aligned}
$$

(iii) *функцията $\mathcal{H}(t, \widetilde{x}(t), p(t), \lambda)$ е непрекъсната в $[\widetilde{t}_0, \widetilde{t}_1]$ и удовлетворява съотношението*

$$
\begin{aligned}
&\mathcal{H}(\widetilde{t}_0, \widetilde{x}(\widetilde{t}_0), p(\widetilde{t}_0), \lambda) = 0,\\
&\mathcal{H}(\widetilde{t}_1, \widetilde{x}(\widetilde{t}_1), p(\widetilde{t}_1), \lambda) = b'_{1_t}(\widetilde{t}_1, \widetilde{x}(\widetilde{t}_1)) \cdot \ell_1.
\end{aligned}
$$

Ще завършим с един елементарен пример, с който ще илюстрираме Принципа за максимума.

Пример 9.1. Ще приложим Принципа за максимума във формата на Хамилтън към екстремалната задача

$$\int_0^1 x(t)\,dt \longrightarrow \inf, \quad \dot{x}(t) = u(t), \quad |u(t)| \le 1, \quad x(0) = 0.$$

Тук времевият интервал е фиксиран — $[0,1]$. Подинтегралната функция е $f(t,x,u) := x$, дясната част на диференциалното уравнение в условието на задачата е $\varphi(t,x,u) := u$, $U := [-1,1]$, като $m = n = 1$. Граничните условия се задават с изображенията $b_0(t,x) := (t,x)$ и $b_1(t,x) := t-1$, като $s_0 = 2$ и $s_1 = 1$. Функцията на Понтрягин има вида $H = pu - \lambda x$.

От Теорема 9.3(i) следва, че съществуват $\lambda \ge 0$, $\ell_0 := (\alpha,\beta) \in \mathbb{R}^2$, $\ell_1 \in \mathbb{R}$ и $p \in AC[0,1]$, не едновременно 0, такива, че удовлетворяват спрегнатото уравнение е $\dot{p} = \lambda$ с гранични условия $p(0) = \beta$ и $p(1) = 0$. Да споменем, че граничните условия в Теорема 9.3(iii) сега имат вида $\mathcal{H}(0,\widetilde{x}(0),p(0),\lambda) = -\alpha$ и $\mathcal{H}(1,\widetilde{x}(1),p(1),\lambda) = \ell_1$. Те, както и $p(0) = \beta$ не носят информация за оптималния процес, както отбелязахме в коментарите по-горе.

Решавайки спрегнатото уравнение, получаваме $p(t) = \lambda t + c$, където $c \in \mathbb{R}$. Граничното условие $p(1) = 0$ влече $p(t) = \lambda(t-1)$.

Ако допуснем, че $\lambda = 0$, получаваме, че всички останали множители на Лагранж са също равни на 0. Следователно $\lambda > 0$ и без ограничение можем да считаме, че $\lambda = 1$. Така имаме $p(t) = t-1$ и $H = (t-1)u - x$. Функцията H достига своя максимум по $u \in [-1,1]$ в $\widetilde{u}(t) := -1$ за всяка функция x, като $\mathcal{H} = 1-t-x$. Решението на диференциалното уравнение $\dot{x} = \widetilde{u}$ с гранично условие $x(0) = 0$ е $\widetilde{x}(t) := -t$.

Така получихме, че единственият допустим процес, който удовлетворява Принципа за максимума, е $\widetilde{x}(t) = -t$, $\widetilde{u}(t) = -1$. Остава да се покаже, че той действително е решение на задачата. Нека $(x(t),u(t))$ е друг допустим процес. Неравенството

$$\int_0^1 x(t)\,dt \ge \int_0^1 \widetilde{x}(t)\,dt$$

се свежда до

$$\int_0^1 (t + x(t))\,dt \ge 0. \tag{9.17}$$

Като вземем предвид, че $\dot{x}(t) = u(t) \ge -1$ и $x(0) = 0$, получаваме

$$x(t) = \int_0^t \dot{x}(\tau)\,d\tau \ge -t,$$

откъдето следва (9.17).

Задачи

1. Изведете от Принципа за максимума уравнението на Ойлер-Лагранж за силен локален минимум за:

   (а) задачата със закрепени краища (Теорема 3.3),

   (б) задачата със свободен край (вж. Зад. 1 в § 3),

   (в) задачата на Болца (вж. Зад. 2 в § 3).

2. Като използвате формулирания тук Принцип за максимума за задачата (9.1), изведете принцип за максимума за минимизиране на функционал от вида $\psi(t_1, x(t_1))$ при условията $(*)$-$(***)$.

   Упътване: Представете $\psi(t_1, x(t_1))$ в интегрален вид, като направите подходящи предположения за гладкостта на ψ.

3. Приложете Принципа за максимума към т.нар. *елементарна задача на оптималното управление*:

$$\left|\begin{array}{l} T \longrightarrow \inf \\ \ddot{x}(t) = u(t), \quad t \in [0,T], \\ |u(t)| \le 1, \quad t \in [0,T], \\ x(0) = \xi_1, \quad \dot{x}(0) = \xi_2, \\ x(T) = \dot{x}(T) = 0. \end{array}\right.$$

   Упътване: Представете $T \longrightarrow \inf$ в интегрален вид и въведете допълнителна променлива $y := \dot{x}$.

4. Решете екстремалната задача от Пример 9.1, като използвате елементарната формула

$$x(t) = \int_0^t \dot{x}(\tau)\, d\tau = \int_0^t u(\tau)\, d\tau.$$

10 Доказателство на Принципа за максимума

Ще изведем Принципа на Понтрягин за максимума от теоремата за смесените задачи в § 8. Ще изложим подхода на Дубовицки и Милютин. Той се основава на смяна на времето.

Нека $(\widetilde{x}(t), \widetilde{u}(t), [\widetilde{t}_0, \widetilde{t}_1])$ е оптимален процес за задачата (9.1). Нека $\widetilde{v}(\eta)$ е неотрицателна, ограничена и измерима функция върху $[0,1]$ такава, че

$$\int_0^1 \widetilde{v}(\eta)\, d\eta = \widetilde{t}_1 - \widetilde{t}_0.$$

Полагаме

$$\widetilde{t}(\tau) := \widetilde{t}_0 + \int_0^\tau \widetilde{v}(\eta)\, d\eta, \quad \tau \in [0,1].$$

Тогава имаме

$$\widetilde{t}(0) = \widetilde{t}_0, \quad \widetilde{t}(1) = \widetilde{t}_1. \tag{10.1}$$

Щом $\widetilde{v}(\eta)$ е ограничена измерима функция върху $[0,1]$, то тя е сумируема върху $[0,1]$ и следователно функцията $\widetilde{t}(\tau)$ е абсолютно непрекъсната. Като вземем предвид, че $\widetilde{v}(\eta)$ е неотрицателна, получаваме още, че $\widetilde{t}(\tau)$ е и монотонно растяща в $[0,1]$.

Въвеждаме още множеството

$$\Delta(\widetilde{v}) := \{\tau \in [0,1] : \widetilde{v}(\tau) > 0\}.$$

Нека $\widetilde{w} : [0,1] \to U$ е измерима функция такава, че

$$\widetilde{w}(\tau) = \widetilde{u}(\widetilde{t}(\tau)) \quad \text{п.н. в} \quad \Delta(\widetilde{v}). \tag{10.2}$$

Тъй като $\widetilde{u}(t)$ е ограничена, то $\widetilde{w}$ е ограничена върху $\Delta(\widetilde{v})$.

Разглеждаме следната екстремална задача

(10.3)

$$\left|\begin{array}{ll}\displaystyle\int_0^1 v(\tau) f(t(\tau), y(\tau), \widetilde{w}(\tau))\, d\tau \longrightarrow \inf & \\ t'(\tau) = v(\tau), \quad y'(\tau) = v(\tau)\, \varphi(t(\tau), y(\tau), \widetilde{w}(\tau)), \quad \tau \in [0,1], & (*) \\ v(\tau) \geq 0, \quad \tau \in [0,1], & (**) \\ b_0(t(0), y(0)) = 0, \quad b_1(t(1), y(1)) = 0. & (***)\end{array}\right.$$

В тази задача фазовите траектории се представят във вида $(t(\tau), y(\tau))$, където $t \in AC[0,1]$ и $y \in AC([0,1], \mathbb{R}^n)$, а управленията $v(\tau)$ са ограничени, измерими, неотрицателни и реално-значни функции върху $[0,1]$, всяка анулираща се върху множество от вида

$$\Delta_k := \{\tau \in [0,1] : \left|\widetilde{w}(\tau)\right| \geq k\}$$

за някое $k \in \mathbb{N}_0$, изобщо различно за различните функции $v(\tau)$. Съвкупността от такива функции $v(\tau)$ ще означаваме с $\mathcal{V}$. Разбира се, $(*)$ и $(**)$ в (10.3) се искат изпълнени п.н. в $[0,1]$. Процесите в задача (10.3) ще означаваме с $(t(\tau), y(\tau), v(\tau))$.

Преимуществото на екстремалната задача (10.3) е в това, че тя е линейна относно управлението $v(\tau)$ и интервалът по времето е фиксиран.

Твърдение 10.1. *Нека* $(\widetilde{x}(t), \widetilde{u}(t), [\widetilde{t}_0, \widetilde{t}_1])$ *е оптимален процес за задачата* (9.1). *Полагаме* $\widetilde{y}(\tau) := \widetilde{x}(\widetilde{t}(\tau))$. *Тогава* $(\widetilde{t}(\tau), \widetilde{y}(\tau), \widetilde{v}(\tau))$ *е оптимален процес за задачата* (10.3).

За да се подготвим за доказателството на това твърдение, ще разгледаме по-подробно въведената по-горе смяна на променливата по времето.

Нека $v : [0,1] \to \mathbb{R}$ е неотрицателна, ограничена и измерима функция. Полагаме

$$\begin{aligned} t(\tau) &:= c + \int_0^\tau v(\eta)\, d\eta, \quad \tau \in [0,1], \\ \Delta(v) &:= \{\tau \in [0,1] : v(\tau) > 0\}. \end{aligned}$$

Тук $c \in \mathbb{R}$. Тогава функцията $t(\tau)$ е абсолютно непрекъсната и монотонно растяща. Ако $v(\tau) = 0$ п.н. върху даден интервал, то $t(\tau)$ е константа върху него; и обратно, ако $t(\tau)$ е константа върху даден интервал, то $v(\tau) = 0$ п.н. върху него. Да означим тези интервали с I_k, $k = 1, 2, \dots$, взаимно непресичащи се. Очевидно те са изброимо много (включително краен брой). Да означим стойността на $t(\tau)$ върху I_k с ξ_k. Върху допълнението на $\cup_k I_k$ функцията $t(\tau)$ е обратима. Означаваме обратната ѝ функция с $\tau(t)$. Тя е дефинирана върху множеството $[t(0), t(1)] \setminus \{\xi_1, \xi_2, \dots\}$. Върху $\{\xi_1, \xi_2, \dots\}$ дефинираме $\tau(t)$, като полагаме $\tau(\xi_k) := \min\{\tau \in [0,1] : t(\tau) = \xi_k\}$, освен ако $\xi_k = t(1)$, в който случай полагаме $\tau(t(1)) := 1$.

Всъщност, както лесно се убеждаваме, това представяне е валидно и върху дефиниционната област на обратната функция на $t(\tau)$, т.е. имаме

$$\tau(t) = \begin{cases} \min\{\tau \in [0,1] : t(\tau) = t\}, & t \in [t(0), t(1)), \\ 1, & t = t(1). \end{cases} \tag{10.4}$$

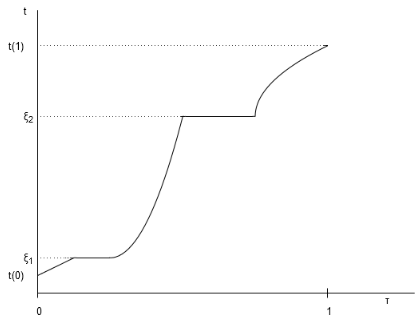
Тази конструкция е представена на фигури 1 и 2 по-долу.

Изразени посредством $\tau(t)$, интервалите I_k се представят така $I_k = [\tau(\xi_k - 0), \tau(\xi_k + 0)]$.

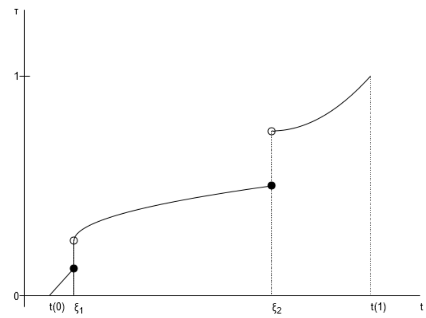
В следващото помощно твърдение ще изясним още връзката между така въведените две функции.

Лема 10.2. *Имаме:*

(a) $t(\tau(\xi)) = \xi$ *за всяко* $\xi \in [t(0), t(1)]$;



Фигура 1: $t = t(\tau)$.



Фигура 2: $\tau = \tau(t)$.

(б) $\tau(t(\eta)) = \eta$ *за почти всяко* $\eta \in \Delta(v)$;

(в) $\tau(t) \in \Delta(v)$ *за почти всяко* $t \in [t(0), t(1)]$.

Доказателство. (а) Това твърдение следва непосредствено от дефиницията на функцията $\tau(t)$ и непрекъснатостта на $t(\tau)$. Действително, ако $\xi \in [t(0), t(1))$ и $\tau(\xi) = \eta$, то $t(\eta) = \xi$. За $\xi = t(1)$ твъдението е съвсем тривиално.

(б) Ако $t(\eta) \neq \xi_k$ за всяко $k \in \mathbb{N}$, т.е. $\eta \notin I_k$ за всяко $k \in \mathbb{N}$, то от дефиницията на $\tau(t)$ следва, че $\tau(t(\eta)) = \eta$. От друга страна върху всеки от интервалите I_k функцията $t(\tau)$ е константа и следователно множеството $\Delta(v) \cap I_k$ има мярка 0 за всяко k. Тогава и тяхното обединение $\Delta(v) \cap (\cup_k I_k)$ също има мярка 0.

(в) Предвид (б) е достатъчно да покажем, че множеството $t(\Delta(v))$ е измеримо и мярката му е равна на дължината на интервала $[t(0), t(1)]$. След като $t(\tau)$ е монотонно растяща абсолютно непрекъсната функция, то за всяко измеримо множество Δ имаме, че $t(\Delta)$ също е измеримо и за лебеговата му мярка $m(t(\Delta))$ е в сила формулата

$$m(t(\Delta)) = \int_\Delta t'(\tau)\, d\tau = \int_\Delta v(\tau)\, d\tau.$$

Множеството $\Delta(v)$ е измеримо, защото v е измерима, и от горната формула получаваме

$$m(t(\Delta(v))) = \int_{\Delta(v)} v(\tau)\, d\tau = \int_0^1 v(\tau)\, d\tau = t(1) - t(0).$$

Така е уставено и третото твърдение на лемата. $\square$

Бележка 10.3. Нека за пълнота отбележим, че, ако $\eta = \min\{\tau : t(\tau) = \xi_k\}$, то, както следва от дефиницията на $\tau(t)$, имаме $\tau(t(\eta)) = \eta$. Така дори имаме, че $\tau(t(\eta)) = \eta$ за всяко $\eta \notin \cup_k(\tau(\xi_k - 0), \tau(\xi_k + 0)]$. Да припомним, че $I_k = [\tau(\xi_k - 0), \tau(\xi_k + 0)]$.

Лема 10.4. *(а) Ако* $x \in AC([t(0), t(1)], \mathbb{R}^n)$ *е решение на диференциалното уравнение* (9.1)-(∗) *п.н. върху интервала* $[t(0), t(1)]$*, то* $y(\tau) := x(t(\tau))$ *е абсолютно непрекъсната векторно-значна функция върху* $[0, 1]$*, която е решение на диференцианото уравнение*

$$y'(\tau) = v(\tau)\, \varphi(t(\tau), y(\tau), w(\tau)) \quad \text{п.н. в} \quad [0, 1], \tag{10.5}$$

където $w : [0, 1] \to U$ *е измерима и* $w(\tau) = u(t(\tau))$ *п.н. в* $\Delta(v)$.

(б) Обратно, ако $w : [0, 1] \to U$ *е измерима функция, която е ограничена върху* $\Delta(v)$*, и ако* $y \in AC([0, 1], \mathbb{R}^n)$ *е решение на* (10.5)*, то* $u(t) := w(\tau(t))$ *е допустимо управление за задачата* (9.1)*,* $x(t) := y(\tau(t))$ *е абсолютно непрекъсната векторно-значна функция върху* $[t(0), t(1)]$*, която е решение на диференциалното уравнение* (9.1)-(∗) *п.н. в* $[t(0), t(1)]$.

Доказателство. (а) Функцията $y(\tau)$ е абсолютно непрекъсната понеже $x(t)$ и $t(\tau)$ са такива и $t(\tau)$ е монотонна. Втората част на твърдение (а) се проверява непосредствено посредством правилото за диференциране на съставни абсолютно непрекъснати функции:

$$\begin{aligned} y'(\tau) &= t'(\tau)\, \dot{x}(t(\tau)) = t'(\tau)\, \varphi(t(\tau), x(t(\tau)), u(t(\tau))) \\ &= v(\tau)\, \varphi(t(\tau), y(\tau), w(\tau)). \end{aligned}$$

За да установим последното равенство, използваме и че $v(\tau) = 0$ в $[0, 1] \backslash \Delta(v)$.

(б) От дефиницията на $u(t)$, направените предположения върху w и Лема 10.2(в) се вижда, че координатните ѝ функции са измерими и ограничени и $u(t) \in U$ п.н. в $[t(0), t(1)]$, т.е. че $u(t)$ е допустимо управление за задачата (9.1).

За да докажем втората част на твърдение (б), използваме, че y е абсолютно непрекъсната, откъдето следва представянето

$$x(t) = y(\tau(t)) = y(0) + \int_0^{\tau(t)} y'(\eta)\, d\eta.$$

Както в (а), посредством (10.5), Лема 10.2(б) и факта, че $v(\tau) = 0$ п.н. в $[0, 1] \backslash \Delta(v)$, достигаме до съотношението

$$\begin{aligned} y'(\eta) &= v(\eta)\, \varphi(t(\eta), y(\eta), w(\eta)) \\ &= t'(\eta)\, \varphi(t(\eta), y(\tau(t(\eta))), w(\tau(t(\eta)))) \\ &= t'(\eta)\, \varphi(t(\eta), x(t(\eta)), u(t(\eta))). \end{aligned}$$

Сега от теоремата за смяна на променливата в лебеговия интеграл и Лема 10.2(а) следва, че

$$x(t) = y(0) + \int_0^t \varphi(\xi, x(\xi), u(\xi))\, d\xi.$$

Оттук следва, че $x(t)$ е абсолютно непрекъсната върху $[t(0), t(1)]$ и $\dot{x}(t) = \varphi(t, x(t), u(t))$ п.н. в $[t(0), t(1)]$. □

Доказателство на Твърдение 10.1. Лема 10.4(а) и условията (10.1) показват, че процесът $(\widetilde{t}(\tau), \widetilde{y}(\tau), \widetilde{v}(\tau))$ е допустим за задачата (10.3).

Нека $(t(\tau), y(\tau), v(\tau))$ е друг допустим процес за задачата (10.3) такъв, че

$$|t(\tau) - \widetilde{t}(\tau)| < \delta, \quad |y(\tau) - \widetilde{y}(\tau)| < \delta, \quad \tau \in [0,1], \tag{10.6}$$

за някое $\delta > 0$.

След като $v(\tau)$ се анулира върху Δ_k за някое $k \in \mathbb{N}_0$, то $\widetilde{w}(\tau)$ е ограничена върху $\Delta(v)$. Да положим $x(t) := y(\tau(t))$ и $u(t) := \widetilde{w}(\tau(t))$. От Леми 10.4(б) и 10.2(в) следва, че $(x(t), u(t), [t(0), t(1)])$ е допустим процес за задачата (9.1). За неговия интервал по t получаваме непосредствено от (10.6)

$$|t(0) - \widetilde{t}_0| < \delta, \quad |t(1) - \widetilde{t}_1| < \delta.$$

Ще оценим разстоянието между фазовите траектории $x(t)$ и $\widetilde{x}(t)$. Първо ще докажем, че $x(t(\tau)) = y(\tau)$ за $\tau \in [0,1]$. От Лема 10.2(б) следва, че $x(t(\tau)) = y(\tau(t(\tau))) = y(\tau)$ за почти всяко $\tau \in \Delta(v)$. Тъй като $x(t)$, $t(\tau)$ и $y(\tau)$ са непрекъснати, то $x(t(\tau)) = y(\tau)$ за всяко $\tau \in \overline{\Delta(v)}$. Ако $\overline{\Delta(v)} \neq [0,1]$, разглеждаме множеството $(0,1)\backslash\overline{\Delta(v)}$. То е отворено. Следователно се представя като обединение на най-много изброимо много два по два непресичащи се отворени интервали. Върху всеки един от тях $v(\tau) = 0$ п.н. Следователно $t(\tau)$ и $y(\tau)$ са константи върху всеки от тях. В краищата на тези интервали $x(t(\tau))$ и $y(\tau)$ са равни. Следователно $x(t(\tau)) = y(\tau)$ за всяко $\tau \in (0,1)\backslash\overline{\Delta(v)}$. Така установихме, че $x(t(\tau)) = y(\tau)$ за $\tau \in (0,1)$. По непрекъснатост това равенство е налице и в краищата на $[0,1]$.

За $t \in [t(0), t(1)] \cap [\widetilde{t}_0, \widetilde{t}_1]$ имаме

$$\begin{aligned}
|x(t) - \widetilde{x}(t)| &= |x(t(\tau)) - \widetilde{x}(t(\tau))| \\
&\leq |x(t(\tau)) - \widetilde{x}(\widetilde{t}(\tau))| + |\widetilde{x}(\widetilde{t}(\tau)) - \widetilde{x}(t(\tau))| \\
&\leq |y(\tau) - \widetilde{y}(\tau)| + \left| \int_{t(\tau)}^{\widetilde{t}(\tau)} |\dot{\widetilde{x}}(\theta)|\, d\theta \right|.
\end{aligned}$$

Оттук като вземем предвид (10.6), както и че $\dot{\widetilde{x}}(\theta) = \varphi(\theta, \widetilde{x}(\theta), \widetilde{u}(\theta))$ (вж. (9.1)-(*)), а $\varphi(t, x, u)$ е ограничена върху ограничените подмножества на V и изображенията $\widetilde{x}(\theta)$ и $\widetilde{u}(\theta)$ са ограничени върху $[\widetilde{t}_0, \widetilde{t}_1]$, получваме, че

$$|x(t) - \widetilde{x}(t)| < \delta + C\,\delta, \quad t \in [t(0), t(1)] \cap [\widetilde{t}_0, \widetilde{t}_1],$$

с някоя положителна константа C.

С разсъждения съвсем аналогични на тези в доказателството на Лема 10.4(б), като използваме оптималността на процеса $(\widetilde{x}(t), \widetilde{u}(t), [\widetilde{t}_0, \widetilde{t}_1])$, получаваме за δ достатъчно малко

$$\begin{aligned}\int_0^1 v(\tau)\, f(t(\tau), y(\tau), \widetilde{w}(\tau))\, d\tau &= \int_{t(0)}^{t(1)} f(t, x(t), u(t))\, dt \\ &\geq \int_{\widetilde{t}_0}^{\widetilde{t}_1} f(t, \widetilde{x}(t), \widetilde{u}(t))\, dt \\ &= \int_0^1 \widetilde{v}(\tau)\, f(\widetilde{t}(\tau), \widetilde{y}(\tau), \widetilde{w}(\tau))\, d\tau,\end{aligned}$$

което показва, че процесът $(\widetilde{t}(\tau), \widetilde{y}(\tau), \widetilde{v}(\tau))$ е оптимален за задачата (10.3). □

Сега ще докажем Принципа за максимума за задачата (10.3). Използваме въведените в предходната глава означения.

Твърдение 10.5. *Нека са изпълнение предположенията, направени в Теорема 9.3. Ако* $(\widetilde{t}(\tau), \widetilde{y}(\tau), \widetilde{v}(\tau))$ *е оптимален процес за задачата* (10.3), *то съществуват* $\lambda \geq 0$, $\ell_0 \in \mathbb{R}^{s_0}$, $\ell_1 \in \mathbb{R}^{s_1}$, $q \in AC([0,1], \mathbb{R}^n)$ *и* $s \in AC[0,1]$, *не всичките нула, такива, че*

(i′) $q'(\tau) = -\widetilde{v}(\tau)\, H'_x(\widetilde{t}(\tau), \widetilde{y}(\tau), \widetilde{w}(\tau), q(\tau), \lambda)$ *п.н. в* $[0,1]$,
$q(i) = (-1)^i b'_{i_x}(\widetilde{t}(i), \widetilde{y}(i))^* \ell_i, \quad i = 0, 1;$

(i″) $s'(\tau) = -\widetilde{v}(\tau)\, H'_t(\widetilde{t}(\tau), \widetilde{y}(\tau), \widetilde{w}(\tau), q(\tau), \lambda)$ *п.н. в* $[0,1]$,
$s(i) = (-1)^i b'_{i_t}(\widetilde{t}(i), \widetilde{y}(i)) \cdot \ell_i, \quad i = 0, 1;$

(ii) $$H(\widetilde{t}(\tau), \widetilde{y}(\tau), \widetilde{w}(\tau), q(\tau), \lambda) + s(\tau) \begin{cases} = 0 & \textit{п.н. в } \Delta(\widetilde{v}), \\ \leq 0 & \textit{п.н. в } [0,1] \backslash \Delta(\widetilde{v}). \end{cases}$$

Доказателство. Ще приложим теоремата за смесени задачи към (10.3). Полагаме:

$$f_0(t, y, v) := \int_0^1 v(\tau) f(t(\tau), y(\tau), \widetilde{w}(\tau))\, d\tau,$$

$$X := AC[0,1] \times AC([0,1], \mathbb{R}^n), \quad Z := L[0,1] \times L([0,1], \mathbb{R}^n) \times \mathbb{R}^{s_0} \times \mathbb{R}^{s_1},$$

$$\Phi_1(t, y, v) := t' - v, \quad \Phi_1 : X \times \mathcal{V} \to L[0,1],$$

$$\Phi_2(t, y, v) := y' - v\, \varphi(t, y, \widetilde{w}), \quad \Phi_2 : X \times \mathcal{V} \to L([0,1], \mathbb{R}^n),$$

$$B_i(t, y, v) := b_i(t(i), y(i)), \quad B_i : X \times \mathcal{V} \to \mathbb{R}^{s_i}, \quad i = 0, 1,$$

$$F(t, y, v) := (\Phi_1, \Phi_2, B_1, B_2), \quad F : X \times \mathcal{V} \to Z.$$

В тези означения задача (10.3) придобива вида

$$\left| \begin{aligned} & f_0(t, y, v) \longrightarrow \inf \\ & F(t, y, v) = 0, \\ & v \in \mathcal{V}. \end{aligned} \right. \tag{10.7}$$

Ще покажем, че тя удовлетворява условията на Теорема 8.2.

Пространството $AC([0,1],\mathbb{R}^n)$ е банахово с норма[1]

$$\|y\|_{W_1^1} := \|y\|_{L([0,1],\mathbb{R}^n)} + \|y'\|_{L([0,1],\mathbb{R}^n)}.$$

(а) За всяко фиксирано $v \in \mathcal{V}$ изображенията f_0 и F са непрекъснато диференцируеми по Фреше в X, като

$$\begin{aligned}
f'_{0_{(t,y)}}(\widetilde{t},\widetilde{y},v)(t,y) &= \int_0^1 v(\tau) f'_t(\widetilde{t}(\tau),\widetilde{y}(\tau),\widetilde{w}(\tau))\, t(\tau)\, d\tau \\
&\qquad + \int_0^1 v(\tau) f'_x(\widetilde{t}(\tau),\widetilde{y}(\tau),\widetilde{w}(\tau)) \cdot y(\tau)\, d\tau, \\
\Phi'_{1_{(t,y)}}(\widetilde{t},\widetilde{y},v)(t,y) &= t', \\
\Phi'_{2_{(t,y)}}(\widetilde{t},\widetilde{y},v)(t,y) &= y' - v\, t\, \varphi'_t(\widetilde{t},\widetilde{y},\widetilde{w}) - v\, \varphi'_x(\widetilde{t},\widetilde{y},\widetilde{w})\, y, \\
B'_{i_{(t,y)}}(\widetilde{t},\widetilde{y},v)(t,y) &= b'_{i_t}(\widetilde{t}(i),\widetilde{y}(i))\, t(i) + b'_{i_x}(\widetilde{t}(i),\widetilde{y}(i))\, y(i), \quad i = 0,1.
\end{aligned}$$

Това следва от твърдения, изведени в 2.3 или се вижда аналогично.

(б) Изображенията f_0 и F удовлетворяват условие (б) в Теорема 8.2, защото са линейни по v и $\mathcal{V}$ е изпъкнало.

(в) Както в доказателството на необходимото условие за задачата на Лагранж, се установява, че $\operatorname{Im} \Phi'_{1_{(t,y)}} = L[0,1]$ и $\operatorname{Im} \Phi'_{2_{(t,y)}} = L([0,1],\mathbb{R}^n)$.[2] Също така очевидно $\operatorname{Im} B'_{i_{(t,y)}}$ имат крайна размерност и коразмерност.

И така Теорема 8.2 е приложима към задачата (10.7). Нейната функция на Лагранж има вида

$$\mathcal{L}(t,y,v,\lambda,z^*) := \lambda f_0(t,y,v) + \langle z^*, F(t,y,v)\rangle,$$

където $\lambda \in \mathbb{R}$ и $z^* \in Z^*$.

От Теорема 8.2 следва, че съществуват множители на Лагранж $\lambda \geq 0$ и $z^* \in Z^*$, не и двата равни на 0, такива, че

$$\mathcal{L}'_{(t,y)}(\widetilde{t},\widetilde{y},\widetilde{v},\lambda,z^*) = 0,$$

т.е.

$$\lambda f'_{0_{(t,y)}}(\widetilde{t},\widetilde{y},\widetilde{v})(t,y) + \langle z^*, F'_{(t,y)}(\widetilde{t},\widetilde{y},\widetilde{v})(t,y)\rangle = 0 \qquad \forall t \in AC[0,1] \quad \forall y \in AC([0,1],\mathbb{R}^n). \tag{10.8}$$

Всеки ограничен линеен функционал z^* над Z се представя във вида

$$\langle z^*, z\rangle = \langle t^*, t\rangle + \langle y^*, y\rangle + \langle r_0^*, r_0\rangle + \langle r_1^*, r_1\rangle, \quad z := (t,y,r_0,r_1) \in Z,$$

[1] Това пространство се означава с $W_1^1([0,1],\mathbb{R}^n)$, а при $n=1$ — с $W_1^1[0,1]$.

[2] Използва се един аналог на познатата теорема за съществуване на решение на линейна система ОДУ, касаещ абсолютно непрекъснати функции. Неговото доказателство е съвсем подобно.

където t^*, y^*, r_0^* и r_1^* са ограничени линейни функционали съответно над $L[0,1]$, $L([0,1],\mathbb{R}^n)$, $\mathbb{R}^{s_0}$ и $\mathbb{R}^{s_1}$. По-рано отбелязахме, че ограничените линейни функционали над $\mathbb{R}^{s_i}$ имат вида

$$\langle r_i^*, a\rangle = \ell_i \cdot a, \quad a \in \mathbb{R}^{s_i},$$

за някой $\ell_i \in \mathbb{R}^{s_i}$.

Според теоремата на Рис за представянето на ограничените линейни функционали над $L[0,1]$ и по-общо над $L([0,1],\mathbb{R}^n)$ имаме (вж. (1.3) и (1.6))

$$\langle t^*, t\rangle = \int_0^1 s(\tau)t(\tau)\,d\tau, \quad t \in L[0,1],$$

с някоя функция $s \in L_\infty[0,1]$ и

$$\langle y^*, y\rangle = \int_0^1 q(\tau)\cdot y(\tau)\,d\tau, \quad y \in L([0,1],\mathbb{R}^n),$$

с някоя векторно-значна функция $q \in L_\infty([0,1],\mathbb{R}^n)$.

Като вземем предвид всички тези резултати за предствянето на линейни функционали, както и посочените в (а) формули за производните на Фреше на f_0 и F, от (10.8) получаваме, че съществуват множители на Лагранж $\lambda \geq 0$, $s \in L_\infty[0,1]$, $q \in L_\infty([0,1],\mathbb{R}^n)$ и $\ell_i \in \mathbb{R}^{s_i}$, $i = 0,1$, не всички равни на 0, такива, че

$$\begin{aligned}
&\int_0^1 \widetilde{v}(\tau) f_t'(\widetilde{t}(\tau),\widetilde{y}(\tau),\widetilde{w}(\tau))\,t(\tau)\,d\tau + \int_0^1 \widetilde{v}(\tau) f_x'(\widetilde{t}(\tau),\widetilde{y}(\tau),\widetilde{w}(\tau))\cdot y(\tau)\,d\tau \\
&+ \int_0^1 s(\tau)t'(\tau)\,d\tau \\
&+ \int_0^1 q(\tau)\cdot\big[y'(\tau) - \widetilde{v}(\tau)\,t(\tau)\,\varphi_t'(\widetilde{t}(\tau),\widetilde{y}(\tau),\widetilde{w}(\tau)) \\
&\qquad\qquad - \widetilde{v}(\tau)\,\varphi_x'(\widetilde{t}(\tau),\widetilde{y}(\tau),\widetilde{w}(\tau))\,y(\tau)\big]\,d\tau \\
&+ \ell_0\cdot\big[b_{0_t}'(\widetilde{t}(0),\widetilde{y}(0))\,t(0) + b_{0_x}'(\widetilde{t}(0),\widetilde{y}(0))\,y(0)\big] \\
&+ \ell_1\cdot\big[b_{1_t}'(\widetilde{t}(1),\widetilde{y}(1))\,t(1) + b_{1_x}'(\widetilde{t}(1),\widetilde{y}(1))\,y(1)\big] = 0 \\
&\qquad\qquad\qquad\qquad \forall t \in AC[0,1] \quad \forall y \in AC([0,1],\mathbb{R}^n).
\end{aligned}$$

Оттук, като веднъж фиксираме $t = 0$ и оставим y произволно, а после обратно — фиксираме $y = 0$ и оставим t произволно, получаваме съответно следните две твърдения:

$$\begin{aligned}
-\int_0^1 &\widetilde{v}(\tau)H_x'(\tau)\cdot y(\tau)\,d\tau + \int_0^1 q(\tau)\cdot y'(\tau)\,d\tau \\
&+ \ell_0\cdot\big(b_{0_x}'(\widetilde{t}(0),\widetilde{y}(0))\,y(0)\big) \\
&+ \ell_1\cdot\big(b_{1_x}'(\widetilde{t}(1),\widetilde{y}(1))\,y(1)\big) = 0 \quad \forall y \in AC([0,1],\mathbb{R}^n),
\end{aligned} \tag{10.9}$$

където за краткост $H'_x(\tau) := H'_x(\widetilde{t}(\tau), \widetilde{y}(\tau), \widetilde{w}(\tau), q(\tau), \lambda)$, и

$$\begin{aligned}(10.10)\qquad -\int_0^1 \widetilde{v}(\tau) H'_t(\tau)\, t(\tau)\, d\tau + \int_0^1 s(\tau)\, t'(\tau)\, d\tau \\ + \big(\ell_0 \cdot b'_{0_t}(\widetilde{t}(0), \widetilde{y}(0)) t(0)\big) \\ + \big(\ell_1 \cdot b'_{1_t}(\widetilde{t}(1), \widetilde{y}(1)) t(1)\big) = 0 \quad \forall t \in AC[0,1],\end{aligned}$$

където $H'_t(\tau) := H'_t(\widetilde{t}(\tau), \widetilde{y}(\tau), \widetilde{w}(\tau), q(\tau), \lambda)$.

В (10.9) внасяме под знака на диференциала $\widetilde{v}(\tau) H'_x(\tau)$, като функцията под знака на диференциала се анулира в 1, и интегрираме по части. Представяме още $y(1)$ посредством формулата $y(1) = y(0) + \int_0^1 y'(\tau)\, d\tau$. Така получаваме, че за всяко $y \in AC([0,1], \mathbb{R}^n)$ имаме

$$\begin{aligned}\int_0^1 \left(q(\tau) + b'_{1_x}(\widetilde{t}(1), \widetilde{y}(1))^* \ell_1 - \int_\tau^1 \widetilde{v}(\eta) H'_x(\eta)\, d\eta \right) \cdot y'(\tau)\, d\tau \\ + y(0) \cdot \left(b'_{0_x}(\widetilde{t}(0), \widetilde{y}(0))^* \ell_0 + b'_{1_x}(\widetilde{t}(1), \widetilde{y}(1))^* \ell_1 - \int_0^1 \widetilde{v}(\tau) H'_x(\tau)\, d\tau \right) = 0.\end{aligned}$$

Следователно

$$q(\tau) + b'_{1_x}(\widetilde{t}(1), \widetilde{y}(1))^* \ell_1 - \int_\tau^1 \widetilde{v}(\eta) H'_x(\eta)\, d\eta = 0 \quad \text{п.н. в} \quad [0,1]$$

и

$$b'_{0_x}(\widetilde{t}(0), \widetilde{y}(0))^* \ell_0 + b'_{1_x}(\widetilde{t}(1), \widetilde{y}(1))^* \ell_1 - \int_0^1 \widetilde{v}(\tau) H'_x(\tau)\, d\tau = 0.$$

Първото от тези съотношения влече, че $q(\tau)$ съвпада п.н. с абсолютно непрекъсната векторно-значна функция и следователно можем да считаме, че $q(\tau)$ е тази функция. Сега от същото съотношение следва, че $q(\tau)$ удовлетворява диференциалното уравнение в (i$'$). Пак от него с $\tau = 1$ получаваме, че q удовлетворява граничното условие в (i$'$) в 1. Накрая от второто съотношение горе следва, че q удовлетворява и граничното условие в 0. С това твърдение (i$'$) е доказано.

Твърдение (i$''$) се установява аналогично от (10.10).

По-нататък според второто твърдение на Теорема 8.2, съотношение (8.3), имаме

$$\mathcal{L}(\widetilde{t}, \widetilde{y}, \widetilde{v}, \lambda, s, q, \ell_0, \ell_1) \le \mathcal{L}(\widetilde{t}, \widetilde{y}, v, \lambda, s, q, \ell_0, \ell_1), \quad v \in \mathcal{V}.$$

Това влече

$$(10.11)\qquad \int_0^1 (v(\tau) - \widetilde{v}(\tau))(H(\tau) + s(\tau))\, d\tau \le 0, \quad v \in \mathcal{V}.$$

Ако $H(\tau)+s(\tau)>0$ върху подмножество S на $\Delta(\widetilde{v})$ с положителна мярка, то с

$$v(\tau) := \begin{cases} 2\widetilde{v}(\tau), & \tau \in S, \\ \widetilde{v}(\tau), & \tau \in [0,1]\backslash S, \end{cases}$$

получаваме

$$\int_0^1 (v(\tau)-\widetilde{v}(\tau))(H(\tau)+s(\tau))\,d\tau = 2\int_S \widetilde{v}(\tau)(H(\tau)+s(\tau))\,d\tau > 0,$$

което противоречи на (10.11). Следователно $H(\tau)+s(\tau)\le 0$ п.н. в $\Delta(\widetilde{v})$.

Да допуснем сега, че $H(\tau)+s(\tau)<0$ върху подмножество T на $\Delta(\widetilde{v})$ с положителна мярка. Тогава с

$$v(\tau) := \begin{cases} \dfrac{\widetilde{v}(\tau)}{2}, & \tau \in T, \\ \widetilde{v}(\tau), & \tau \in [0,1]\backslash T, \end{cases}$$

получаваме

$$\int_0^1 (v(\tau)-\widetilde{v}(\tau))(H(\tau)+s(\tau))\,d\tau = -\frac{1}{2}\int_T \widetilde{v}(\tau)(H(\tau)+s(\tau))\,d\tau > 0,$$

което отново противоречи на (10.11). Следователно $H(\tau)+s(\tau)=0$ п.н. в $\Delta(\widetilde{v})$.

Така първата част на (ii) е установена.

Втората се установява аналогично, като вземем предвид, че $\widetilde{v}(t)=0$ п.н. върху $[0,1]\backslash\Delta(\widetilde{v})$. □

Готови сме да преминем към доказателство на Принципа за максимума за задачата (9.1).

Доказателство на Теорема 9.3. Нека $(\widetilde{x}(t),\widetilde{u}(t),[\widetilde{t}_0,\widetilde{t}_1])$ е оптимален процес за задачата (9.1). Тогава според Твърдение 10.1 $(\widetilde{t}(\tau),\widetilde{y}(\tau),\widetilde{v}(\tau))$ е оптимален процес за задачата (10.3). Да припомним, $\widetilde{y}(\tau) := \widetilde{x}(\widetilde{t}(\tau))$. Прилагаме Твърдение 10.5.

Полагаме

$$\widetilde{\tau}(t) := \begin{cases} \min\{\tau\in[0,1] : \widetilde{t}(\tau)=t\}, & t\in[\widetilde{t}_0,\widetilde{t}_1), \\ 1, & t=\widetilde{t}_1, \end{cases}$$

т.е. $\widetilde{\tau}(t)$ отговаря на $\widetilde{t}(\tau)$, както $\tau(t)$ на $t(\tau)$. Полагаме още $p(t) := q(\widetilde{\tau}(t))$. От Твърдение 10.5(i′) веднага следва, че $p(t)$ удовлетворява граничните условия в Теорема 9.3(i).

Измеримата функция $\widetilde{w}(\tau)$ е ограничена върху $\Delta(\widetilde{v})$. От Лема 10.2(в) следва, че $\widetilde{\tau}(t)\in\Delta(\widetilde{v})$ п.н. в $[\widetilde{t}_0,\widetilde{t}_1]$, откъдето, прилагайки Лема 10.2(а) получаваме $\widetilde{w}(\widetilde{\tau}(t)) = \widetilde{u}(\widetilde{t}(\widetilde{\tau}(t))) = \widetilde{u}(t)$ за почти всяко $t\in[\widetilde{t}_0,\widetilde{t}_1]$. След това с разсъждения съвсем сходни на тези в доказателството на Лема 10.4(б), установяваме, че $p \in AC([\widetilde{t}_0, \widetilde{t}_1], \mathbb{R}^n)$ и $p(t)$ удовлетворява и диференциалното уравнение в Теорема 9.3(i). С това твърдение (i) на Теорема 9.3 е доказано. Също така с разсъждения аналогични на използваните в доказателството на Твърдение 10.1, установяваме, че $p(\widetilde{t}(\tau)) = q(\tau)$, $\tau \in [0,1]$.

Аналогично, за $r(t) := s(\widetilde{\tau}(t))$ получаваме посредством Твърдение 10.5(i″), че $r \in AC[\widetilde{t}_0, \widetilde{t}_1]$,

$$\dot{r}(t) = -H'_t(t, \widetilde{x}(t), \widetilde{u}(t), p(t), \lambda) \quad \text{п.н. в} \quad [\widetilde{t}_0, \widetilde{t}_1] \tag{10.12}$$

$$r(\widetilde{t}_i) = (-1)^i b'_{i_t}(\widetilde{t}_i, \widetilde{x}(\widetilde{t}_i)) \cdot \ell_i, \quad i = 0, 1, \tag{10.13}$$

и $r(\widetilde{t}(\tau)) = s(\tau)$, $\tau \in [0,1]$.

За останалата част на доказателството ще подберем функциите $\widetilde{v}$ и $\widetilde{w}$ така, че освен да изпълняват условията (10.1) и (10.2), да имат още две свойства.

Нека $\widetilde{v}$ е такава, че се анулира върху два по два непресичащите се интервали I_k, $k = 1, 2, \ldots$, от вида $(\cdot,\cdot]$ такива, че множеството $\widetilde{t}(\cup_k I_k)$ е навсякъде гъсто в $[\widetilde{t}_0, \widetilde{t}_1]$. По-точно фиксираме $\widetilde{v}$ така, че $\widetilde{t}(\tau) = \xi_k$, $\tau \in I_k$, и множеството $\{\xi_k\}$ е навсякъде гъсто в $[\widetilde{t}_0, \widetilde{t}_1]$.[1]

Нека изброимото множество $\{u_i\}$ е навсякъде гъсто в U. Разлагаме всяко I_k в обединение на непресичащи се два по два интервали I_{ki}, $i = 1, 2 \ldots$, също от вида $(\cdot,\cdot]$. Дефинираме $\widetilde{w}(\tau)$ върху $\cup_k I_k \subseteq [0,1] \backslash \Delta(\widetilde{v})$, като полагаме $\widetilde{w}(\tau) := u_i$ п.н. в I_{ki}.

От неравенството в Твърдение 10.5(ii) следва, че

$$H(\widetilde{t}(\tau), \widetilde{y}(\tau), \widetilde{w}(\tau), q(\tau), \lambda) + s(\tau) \le 0 \quad \text{п.н. в} \quad \bigcup_k I_k,$$

откъдето получаваме

$$H(\widetilde{t}(\tau), \widetilde{x}(\widetilde{t}(\tau)), \widetilde{w}(\tau), p(\widetilde{t}(\tau)), \lambda) + r(\widetilde{t}(\tau)) \le 0 \quad \text{п.н. в} \quad \bigcup_k I_k.$$

Оттук, като вземем предвид, че $\widetilde{t}(\tau) = \xi_k$ в I_k и $\widetilde{w}(\tau) = u_i$ п.н. в I_{ki}, получаваме

$$H(\xi_k, \widetilde{x}(\xi_k), u_i, p(\xi_k), \lambda) + r(\xi_k) \le 0 \quad \forall k \quad \forall i.$$

Като използваме, че $\{\xi_k\}$ е навсякъде гъсто в $[\widetilde{t}_0, \widetilde{t}_1]$, $\{u_i\}$ е навсякъде гъсто в U и че $H(t, \widetilde{x}(t), u, p(t), \lambda)$ е непрекъсната по (t, u), заключаваме, че

$$H(t, \widetilde{x}(t), u, p(t), \lambda) + r(t) \le 0 \quad \forall t \in [\widetilde{t}_0, \widetilde{t}_1] \quad \forall u \in U. \tag{10.14}$$

[1] След края на доказателството е даден пример на такава функция.

От друга страна от равенството в Твърдение 10.5(ii) и Лема 10.2, (а) и (в), следва, че

$$H(t,\widetilde{x}(t),\widetilde{u}(t),p(t),\lambda)+r(t)=0 \quad \text{п.н. в} \quad [\widetilde{t}_0,\widetilde{t}_1]. \tag{10.15}$$

Сега твърдение (ii) на Теорема 9.3 следва от (10.14) и (10.15).

Накрая, като използваме отново, че функцията $H(t,\widetilde{x}(t),u,p(t),\lambda)$ е непрекъсната по (t,u) и изображението $\widetilde{u}(t)$ е ограничено, установяваме, че $\mathcal{H}(t,\widetilde{x}(t),p(t),\lambda)$ е непрекъсната по t. Тъй като $r(t)$ е също непрекъсната, от Теорема 9.3(ii) и (10.15) следва, че $\mathcal{H}(t,\widetilde{x}(t),p(t),\lambda)=-r(t)$ за всяко $t\in[\widetilde{t}_0,\widetilde{t}_1]$. Сега твърдение (iii) на Теорема 9.3 следва от граничните условия (10.13), а формула (9.3) пък следва от (10.12)-(10.13) и това, че $r(t)$ е абсолютно непрекъсната. □

Дефиниция на функцията $\widetilde{v}(\tau)$: Нека $\{\xi_k\}$ е изброимо навсякъде гъсто подмножество на $[\widetilde{t}_0,\widetilde{t}_1]$. Например, можем да вземем множеството на рационалните числа в $[\widetilde{t}_0,\widetilde{t}_1]$. Фиксираме положителни числа β_k, $k=1,2,\dots$, такива, че

$$\sum_{k=1}^{\infty}\beta_k=\frac{1}{2}.$$

Например, можем да вземем $\beta_k=2^{-k-1}$.

Полагаме

$$\tau_k:=\frac{\xi_k-\widetilde{t}_0}{2(\widetilde{t}_1-\widetilde{t}_0)}+\sum_{\xi_i<\xi_k}\beta_i.$$

Дефинираме интервалите $I_k:=(\tau_k,\tau_k+\beta_k]$. Ясно е, че $I_k\subseteq[0,1]$ за всяко k. Също така те са два по два непресичащи се. Действително, ако естествените числа k и ℓ са такива, че $\xi_k<\xi_\ell$, то $\tau_k+\beta_k<\tau_\ell$.

Дефинираме функцията $\widetilde{v}(\tau)$ върху $[0,1]$, като пологаме

$$\widetilde{v}(\tau):=\begin{cases}0, & \tau\in\bigcup_k I_k,\\ 2(\widetilde{t}_1-\widetilde{t}_0) & \text{в противен случай.}\end{cases}$$

Очевидно $\widetilde{v}(\tau)$ е неотрицателна, ограничена и измерима функция. За нейния интеграл намираме непосредствено

$$\int_0^1\widetilde{v}(\tau)\,d\tau=2(\widetilde{t}_1-\widetilde{t}_0)\int_{[0,1]\setminus(\cup_k I_k)}d\tau=2(\widetilde{t}_1-\widetilde{t}_0)\left(1-\sum_{k=1}^{\infty}\beta_k\right)=\widetilde{t}_1-\widetilde{t}_0.$$

Ще покажем, че $\widetilde{t}(\tau) = \xi_k$, ако $\tau \in I_k$. Нека $\tau \in I_k$. Тогава

$$\begin{aligned}
\widetilde{t}(\tau) &:= \widetilde{t}_0 + \int_0^{\tau} \widetilde{v}(\eta)\, d\eta = \widetilde{t}_0 + \int_0^{\tau_k} \widetilde{v}(\eta)\, d\eta \\
&= \widetilde{t}_0 + 2(\widetilde{t}_1 - \widetilde{t}_0) \int_{[0,\tau_k]\setminus(\cup_{\tau_j<\tau_k} I_j)} d\eta \\
&= \widetilde{t}_0 + 2(\widetilde{t}_1 - \widetilde{t}_0)\Bigg(\tau_k - \sum_{\xi_j<\xi_k} \beta_j\Bigg) \\
&= \widetilde{t}_0 + 2(\widetilde{t}_1 - \widetilde{t}_0)\frac{\xi_k - \widetilde{t}_0}{2(\widetilde{t}_1 - \widetilde{t}_0)} = \xi_k.
\end{aligned}$$

Задачи

1. Уверете се във верността на формулите в (а) в доказателството на Твърдение 10.5.

2. Проследете малките модификации в доказателството на Теорема 9.3, с които се установя Теорема 9.8.

11 Задачи с фазови ограничения

Тук ще формулираме Принципа за максимума за един по-общ вид задачи в оптималното управление. В тях са включени т.нар. фазови ограничения.

Нека са дадени изображенията

$$\begin{aligned}
&f : V \to \mathbb{R},\\
&\varphi : V \to \mathbb{R}^n,\\
&b_i : V_i' \to \mathbb{R}^{s_i}, \quad i = 0, 1,\\
&g_i : V_i'' \to \mathbb{R}, \quad i = 0, 1, \ldots, k,
\end{aligned}$$

където $V \subseteq \mathbb{R}^{m+n+1}$, $V_i' \subseteq \mathbb{R}^n$, $i = 0, 1$, и $V_i'' \subseteq \mathbb{R}^{n+1}$. Нека също $U \subseteq \mathbb{R}^m$. Разглеждаме екстремалната задача

$$(11.1) \qquad \left| \begin{array}{ll}
\mathcal{F}(x, u) := \displaystyle\int_{t_0}^{t_1} f(t, x(t), u(t))\, dt \longrightarrow \inf & \\
\dot{x}(t) = \varphi(t, x(t), u(t)), \quad t \in [t_0, t_1], & (*)\\
u(t) \in U, \quad t \in [t_0, t_1], & (**)\\
b_0(x(t_0)) = 0, \quad b_1(x(t_1)) = 0, & (***)\\
g_i(t, x(t)) \le 0, \quad t \in [t_0, t_1],\ i = 1, 2, \ldots, k. & (****)
\end{array} \right.$$

Както в предходните глави 9 и 10, предполагаме, че фазовите траектории $x(t) := (x_1(t), x_2(t), \ldots, x_n(t))$ са векторно-значни функции, чиито координатни функции са абсолютно непрекъснати и удовлетворяват $(*)$ почти навсякъде относно лебеговата мярка; управленията $u(t) := (u_1(t), u_2(t), \ldots, u_m(t))$ са векторно-значни функции с координатни функции, които са ограничени, измерими (отново по Лебег) и удовлетворяват $(**)$ почти навсякъде. За начало ще считаме, че интервалът $[t_0, t_1]$ е *фиксиран.*

Условията $(****)$ се наричат *фазови ограничения.*

Следва Принципът за максимума за тези задачи.

Теорема 11.1 (Принцип за максимума във форма на Хамилтън)**.** *Нека $f, \varphi \in C(V)$, $b_i \in C^1(V_i')$, $i = 0, 1$, $g_i \in C(V_i'')$, $i = 0, 1, \ldots, k$, където $V \subseteq \mathbb{R}^{n+m+1}$, $V_i' \subseteq \mathbb{R}^n$ и $V_i'' \subseteq \mathbb{R}^{n+1}$ са отворени. Нека f, φ и g_i са непрекъснато диференцируеми по $x_1, x_2, \ldots, x_n$. Ако $(\widetilde{x}(t), \widetilde{u}(t))$ е оптимален процес за задачата (11.1), то съществуват $\lambda \ge 0$, $\ell_i \in \mathbb{R}^{s_i}$, $i = 0, 1$, непрекъсната отляво векторно-значна функция с ограничена вариация $p : [t_0, t_1] \to \mathbb{R}^n$ и регулярни неотрицателни борелови мерки*

μ_i, $i = 1, 2, \ldots, k$, върху $[t_0, t_1]$, не всичките нула, такива, че

$$
\begin{aligned}
\text{(i)} \qquad & p(t) = -b_1'(\widetilde{x}(t_1))^* \ell_1 + \int_t^{t_1} H_x'(\xi, \widetilde{x}(\xi), \widetilde{u}(\xi), p(\xi), \lambda)\, d\xi \\
& \qquad\qquad - \sum_{i=1}^{k} \int_t^{t_1} g_{i_x}'(\xi, \widetilde{x}(\xi))\, d\mu_i, \quad t \in (t_0, t_1], \\
& p(t_0) = b_0'(\widetilde{x}(t_0))^* \ell_0; \\
\text{(ii)} \qquad & H(t, \widetilde{x}(t), \widetilde{u}(t), p(t), \lambda) = \mathcal{H}(t, \widetilde{x}(t), p(t), \lambda) \quad \text{п.н. в} \quad [t_0, t_1].
\end{aligned}
$$

Мерките μ_i, $i = 1, 2, \ldots, k$, са с носители съответно множествата

$$T_i := \{t \in [t_0, t_1] : g_i(t, \widetilde{x}(t)) = 0\}.$$

Както в предходните две теми, H е функцията на Понтрягин

$$H(t, x, u, p, \lambda) := p \cdot \varphi(t, x, u) - \lambda f(t, x, u).$$

а функцията

$$\mathcal{H}(t, x, p, \lambda) := \sup_{u \in U} H(t, x, u, p, \lambda)$$

е хамилтонианът на задачата (11.1). Уравнението относно $p(t)$ в първото твърдение на теоремата се нарича *спрегнато*. При отсъствието на фазови ограничения, то се свежда до спрегнатото уравнение в Теорема 9.3. В този случай векторно-значната функция $p(t)$ е абсолютно непрекъсната. В общия случай тя е само функция с ограничена вариация и непрекъсната отляво, като е възможно да има точки на прекъсване.

Забележете, че носителите на мерките μ_i са точно множествата, върху които фазовите ограничения се реализират като равенства.

Теорема 11.1 може да се докаже аналогично на Теорема 9.3, като се сведе към смесена задача с фазови ограничения. За нея съществува аналогично необходимо условие на разгледаното в § 8 (вж. Теорема 1 в § 5.1 в книгата на Йофе и Тихомиров).

В екстремалната задача (11.1) интервалът по времето е фиксиран. Тогава фазовите траектории попадат в определено естествено банахово пространство и не е необходимо да предполагаме гладкост на разглежданите фунции по времето. В противен случай, т.е. при променлив времеви интервал, фазовите траектории нямат това свойство и трябва да разглеждаме времето като фазова променлива.

Задачата с фазови ограничения и свободно време лесно се свежда към такава с фиксирано. Да разгледаме задачата

$$
(11.2)\qquad \left|\begin{array}{ll}
\mathcal{F}(x,u) := \displaystyle\int_{t_0}^{t_1} f(t,x(t),u(t))\,dt \longrightarrow \inf & \\
\dot{x}(t) = \varphi(t,x(t),u(t)), \quad t\in[t_0,t_1], & (*)\\
u(t)\in U, \quad t\in[t_0,t_1], & (**)\\
b_0(t_0,x(t_0)) = 0, \quad b_1(t_1,x(t_1)) = 0, & (***)\\
g_i(t,x(t)) \le 0, \quad t\in[t_0,t_1],\ i=1,2,\dots,k, & (****)
\end{array}\right.
$$

като интервалът по времето $[t_0,t_1]$ е променлив.

Освен направените в Теорема 11.1 предположения, нека още **всички изображения са непрекъснато диференцируеми по** t.

Нека $(t(\tau),y(\tau))$, $\tau\in[0,1]$, е решение на системата

$$
(11.3)\qquad \left|\begin{array}{l}
t'(\tau) = v(\tau)\\
y'(\tau) = v(\tau)\,\varphi(t(\tau),y(\tau),w(\tau)),
\end{array}\right.
$$

където $v\in C[0,1]$ и $w(\tau)$ е ограничена измерима функция върху $[0,1]$. Ако $v(\tau)>0$ върху $[0,1]$, то $t(\tau)$ е строго растяща непрекъснато диференцируема функция. Тогава тя е обратима и нейната обратна $\tau(t)$ също е строго растяща и непрекъснато диференцируема. При това смените на променливата, които се дефинират чрез трансформациите $t=t(\tau)$ и $\tau=\tau(t)$ запазват свойствата: измеримост, ограниченост и абсолютна непрекъснатост. Също така, ако $y(\tau)$ удовлетворява второто диференциално уравнение в (11.3), то $x(t) := y(\tau(t))$ ще удовлетворява (11.1)-($*$) с $u(t) := w(\tau(t))$ п.н. върху интервала $[t(0),t(1)]$. Обратно, ако $x(t)$ е решение на (11.1)-($*$), което отговаря на управлението $u(t)$, върху $[t_0,t_1]$, то функциите $t(\tau) := t_0+(t_1-t_0)\tau$, $y(\tau) := x(t(\tau))$ са решение на (11.3), което отговаря на управленията $v(\tau) := t_1-t_0$ и $w(\tau) := u(t(\tau))$.

По-натък, съвсем аналогично както в предходната глава, се установява, че ако $(\widetilde{x}(t),\widetilde{u}(t),[\widetilde{t}_0,\widetilde{t}_1])$ е оптимален процес за задачата (11.2), то

$$
\begin{aligned}
&\widetilde{t}(\tau) := \widetilde{t}_0 + (\widetilde{t}_1-\widetilde{t}_0)\tau, \quad \widetilde{y}(\tau) := \widetilde{x}(\widetilde{t}(\tau)),\\
&\widetilde{v}(\tau) := \widetilde{t}_1-\widetilde{t}_0, \quad \widetilde{w}(\tau) := \widetilde{u}(\widetilde{t}(\tau))
\end{aligned}
$$

е оптимален процес за задачата

$$
(11.4)\qquad \left|\begin{array}{ll}
\mathcal{F}(x,u) := \displaystyle\int_0^1 v(\tau)\,f(t(\tau),y(\tau),w(\tau))\,d\tau \longrightarrow \inf & \\
t'(\tau) = v(\tau), \quad y'(\tau) = v(\tau)\,\varphi(t(\tau),y(\tau),w(\tau)),\ \tau\in[0,1], & (*)\\
v(\tau)>0, \quad w(\tau)\in U, \quad \tau\in[0,1], & (**)\\
b_0(t(0),y(0)) = 0, \quad b_1(t(1),y(1)) = 0, & (***)\\
g_i(t(\tau),y(\tau)) \le 0, \quad \tau\in[0,1],\ i=1,2,\dots,k. & (****)
\end{array}\right.
$$

Това е задача от типа на (11.1) с фиксирано време, към която Теорема 11.1 е приложима. Нека $\widetilde{H}$ е нейната функция на Понтрягин:

$$\widetilde{H}(t, y, v, w, r, s, \lambda) := r \cdot v\,\varphi(t, y, w) + s\,v - \lambda\, v\, f(t, y, w).$$

Следователно

$$\widetilde{H}(t, y, v, w, r, s, \lambda) = v(H(t, y, w, r, \lambda) + s). \tag{11.5}$$

Да положим още

$$\widetilde{T}_i := \{\tau \in [0,1] : g_i(\widetilde{t}(\tau), \widetilde{y}(\tau)) = 0\}, \quad k = 1, 2, \ldots, k.$$

Нека $(\widetilde{t}(\tau), \widetilde{y}(\tau), \widetilde{v}(\tau), \widetilde{w}(\tau))$ е решение на задачата (11.4). Полагаме

$$\widetilde{H}(\tau) := \widetilde{H}(\widetilde{t}(\tau), \widetilde{y}(\tau), \widetilde{v}(\tau), \widetilde{w}(\tau), r(\tau), s(\tau), \lambda).$$

От Теорема 11.1 следва, че съществуват $\lambda \geq 0$, $\ell_i \in \mathbb{R}^{s_i}$, $i = 0, 1$, непрекъснати отляво векторно-значни функции с ограничена вариация $r : [0,1] \to \mathbb{R}^n$ и $s : [0,1] \to \mathbb{R}$ и регулярни борелови мерки ν_i, $i = 1, 2, \ldots, k$, върху $[0,1]$ с носители съответно множествата $\widetilde{T}_i$, не всичките нула, такива, че

$$r(\tau) = -b'_{1_x}(\widetilde{t}(1), \widetilde{y}(1))^* \ell_1 + \int_\tau^1 \widetilde{H}'_x(\eta)\, d\eta - \sum_{i=1}^k \int_\tau^1 g'_{i_x}(\widetilde{t}(\eta), \widetilde{y}(\eta))\, d\nu_i, \quad \tau \in (0,1], \tag{11.6}$$

$$s(\tau) = -b'_{1_t}(\widetilde{t}(1), \widetilde{y}(1)) \cdot \ell_1 + \int_\tau^1 \widetilde{H}'_t(\eta)\, d\eta - \sum_{i=1}^k \int_\tau^1 g'_{i_t}(\widetilde{t}(\eta), \widetilde{y}(\eta))\, d\nu_i, \quad \tau \in (0,1], \tag{11.7}$$

$$r(0) = b'_{0_x}(\widetilde{t}(0), \widetilde{y}(0))^* \ell_0, \tag{11.8}$$

$$s(0) = b'_{0_t}(\widetilde{t}(0), \widetilde{y}(0)) \cdot \ell_0 \tag{11.9}$$

и

$$\widetilde{H}(\tau) = \max_{u \in U,\, v > 0} \widetilde{H}(\widetilde{t}(\tau), \widetilde{y}(\tau), v, u, r(\tau), s(\tau), \lambda) \quad \text{п.н. в} \quad [0,1]. \tag{11.10}$$

Нека $\widetilde{\tau}(t)$ е обратната функция на $\widetilde{t}(\tau)$, т.е.

$$\widetilde{\tau}(t) := \frac{t - \widetilde{t}_0}{\widetilde{t}_1 - \widetilde{t}_0}.$$

Полагаме $p(t) := r(\widetilde{\tau}(t))$ и $q(t) := s(\widetilde{\tau}(t))$ за $t \in [\widetilde{t}_0, \widetilde{t}_1]$. След като r и s са непрекъснати отляво функции с ограничена вариация върху $[0,1]$, то и

p и q са такива върху $[\widetilde{t}_0, \widetilde{t}_1]$. Нека също така за $i = 1, 2, \ldots, k$ положим $\mu_i(S) := \nu_i\big(\widetilde{\tau}(S)\big)$, където S е борелово подмножество на $[\widetilde{t}_0, \widetilde{t}_1]$. Тогава μ_i е регулярна борелова мярка върху $[\widetilde{t}_0, \widetilde{t}_1]$ с носител множеството

$$T_i := \widetilde{t}(\widetilde{T}_i) = \{t \in [\widetilde{t}_0, \widetilde{t}_1] : g_i(t, \widetilde{x}(t)) = 0\}.$$

От дефиницията на μ_i следва, че за всеки $a, b \in [\widetilde{t}_0, \widetilde{t}_1]$

$$\int_a^b \psi(t)\, d\mu_i = \int_{\widetilde{\tau}(a)}^{\widetilde{\tau}(b)} \psi(\widetilde{t}(\tau))\, d\nu_i \quad \forall \psi \in C[\widetilde{t}_0, \widetilde{t}_1].$$

В така въведените означения, формулите (11.6)-(11.10) приемат следния вид след смяна на променливата $\tau = \widetilde{\tau}(t)$:

$$\begin{aligned} p(t) = -b'_{1_x}(\widetilde{t}_1, \widetilde{x}(\widetilde{t}_1))^* \ell_1 &+ \int_t^{\widetilde{t}_1} H'_x(\xi, \widetilde{x}(\xi), \widetilde{u}(\xi), p(\xi), \lambda)\, d\xi \\ &- \sum_{i=1}^k \int_t^{\widetilde{t}_1} g'_{i_x}(\xi, \widetilde{x}(\xi))\, d\mu_i, \quad t \in (\widetilde{t}_0, \widetilde{t}_1], \end{aligned} \tag{11.11}$$

$$\begin{aligned} q(t) = -b'_{1_t}(\widetilde{t}_1, \widetilde{x}(\widetilde{t}_1)) \cdot \ell_1 &+ \int_t^{\widetilde{t}_1} H'_t(\xi, \widetilde{x}(\xi), \widetilde{u}(\xi), p(\xi), \lambda)\, d\xi \\ &- \sum_{i=1}^k \int_t^{\widetilde{t}_1} g'_{i_t}(\xi, \widetilde{x}(\xi))\, d\mu_i, \quad t \in (\widetilde{t}_0, \widetilde{t}_1], \end{aligned} \tag{11.12}$$

$$p(\widetilde{t}_0) = b'_{0_x}(\widetilde{t}_0), \widetilde{x}(\widetilde{t}_0))^* \ell_0, \tag{11.13}$$

$$q(\widetilde{t}_0) = b'_{0_t}(\widetilde{t}_0, \widetilde{x}(\widetilde{t}_0)) \cdot \ell_0 \tag{11.14}$$

и, като вземем предвид (11.5),

$$\begin{aligned} &\widetilde{v}(t)\big(H(t, \widetilde{x}(t), \widetilde{u}(t), p(t), \lambda) + q(t)\big) \\ &\quad = \max_{u \in U,\, v > 0} v\big(H(t, \widetilde{x}(t), u, p(t), \lambda) + q(t)\big) \quad \text{п.н. в} \quad [\widetilde{t}_0, \widetilde{t}_1]. \end{aligned} \tag{11.15}$$

Последното влече

$$\max_{u \in U} H(t, \widetilde{x}(t), u, p(t), \lambda) = -q(t) \quad \text{п.н. в} \quad [\widetilde{t}_0, \widetilde{t}_1], \tag{11.16}$$

иначе функцията на u и v в дясната страна на (11.15) не би достигнала крайна най-голяма стойност. Тогава от (11.15) следва и

$$H(t, \widetilde{x}(t), \widetilde{u}(t), p(t), \lambda) = -q(t) \quad \text{п.н. в} \quad [\widetilde{t}_0, \widetilde{t}_1]. \tag{11.17}$$

Равенствата (11.16) и (11.17) влекат

$$H(t, \widetilde{x}(t), \widetilde{u}(t), p(t), \lambda) = \mathcal{H}(t, \widetilde{x}(t), p(t), \lambda) \quad \text{п.н. в} \quad [\widetilde{t}_0, \widetilde{t}_1]. \tag{11.18}$$

От (11.12), (11.14) и (11.16) следва, че функцията $\mathcal{H}(t, \widetilde{x}(t), p_1(t), \lambda)$ съвпада почти навсякъде върху $[\widetilde{t}_0, \widetilde{t}_1]$ с непрекъсната отляво функция с ограничена вариация, при това са в сила формулите

$$\begin{aligned}\mathcal{H}(t, \widetilde{x}(t), p(t), \lambda) = b'_{1_t}(\widetilde{t}_1, \widetilde{x}(\widetilde{t}_1)) \cdot \ell_1 - \int_t^{\widetilde{t}_1} H'_t(\xi, \widetilde{x}(\xi), \widetilde{u}(\xi), p(\xi), \lambda)\, d\xi \\ + \sum_{i=1}^k \int_t^{\widetilde{t}_1} g'_{i_t}(\xi, \widetilde{x}(\xi))\, d\mu_i \quad \text{п.н. в} \quad t \in (\widetilde{t}_0, \widetilde{t}_1], \\ \mathcal{H}(\widetilde{t}_0, \widetilde{x}(\widetilde{t}_0), p(\widetilde{t}_0), \lambda) = -b'_{0_t}(\widetilde{t}_0, \widetilde{x}(\widetilde{t}_0)) \cdot \ell_0.\end{aligned}$$

Това е аналог на Теорема 9.3(iii) и формула (9.3).

Формули (11.11), (11.13), (11.18) и последните две съотношения горе съставляват Принципа за максимума за задачата (11.2):

Теорема 11.2 (Принцип за максимума във форма на Хамилтън)**.** *Нека $f, \varphi \in C(V)$, $b_i \in C^1(V'_i)$, $i = 0, 1$, $g_i \in C^1(V''_i)$, $i = 0, 1, \dots, k$, където $V \subseteq \mathbb{R}^{n+m+1}$ и $V'_i, V''_i \subseteq \mathbb{R}^{n+1}$ са отворени. Нека f и φ са непрекъснато диференцируеми по $t, x_1, x_2, \dots, x_n$. Ако $(\widetilde{x}(t), \widetilde{u}(t), [\widetilde{t}_0, \widetilde{t}_1])$ е оптимален процес за задачата (11.2), то съществуват $\lambda \geq 0$, $\ell_i \in \mathbb{R}^{s_i}$, $i = 0, 1$, непрекъсната отляво векторно-значна функция с ограничена вариация $p : [\widetilde{t}_0, \widetilde{t}_1] \to \mathbb{R}^n$ и регулярни борелови мерки μ_i, $i = 1, 2, \dots, k$, върху $[\widetilde{t}_0, \widetilde{t}_1]$, не всичките нула, такива, че*

(i) $$\begin{aligned}p(t) = -b'_{1_x}(\widetilde{t}_1, \widetilde{x}(\widetilde{t}_1))^* \ell_1 + \int_t^{\widetilde{t}_1} H'_x(\xi, \widetilde{x}(\xi), \widetilde{u}(\xi), p(\xi), \lambda)\, d\xi \\ - \sum_{i=1}^k \int_t^{\widetilde{t}_1} g'_{i_x}(\xi, \widetilde{x}(\xi))\, d\mu_i, \quad t \in (\widetilde{t}_0, \widetilde{t}_1], \\ p(\widetilde{t}_0) = b'_{0_x}(\widetilde{t}_0), \widetilde{x}(\widetilde{t}_0))^* \ell_0;\end{aligned}$$

(ii) $H(t, \widetilde{x}(t), \widetilde{u}(t), p_1(t), \lambda) = \mathcal{H}(t, \widetilde{x}(t), p_1(t), \lambda)$ *п.н. в* $[\widetilde{t}_0, \widetilde{t}_1]$.

(iii)
$$\begin{aligned}\mathcal{H}(t, \widetilde{x}(t), p(t), \lambda) = b'_{1_t}(\widetilde{t}_1, \widetilde{x}(\widetilde{t}_1)) \cdot \ell_1 \\ - \int_t^{\widetilde{t}_1} H'_t(\xi, \widetilde{x}(\xi), \widetilde{u}(\xi), p(\xi), \lambda)\, d\xi \\ + \sum_{i=1}^k \int_t^{\widetilde{t}_1} g'_{i_t}(\xi, \widetilde{x}(\xi))\, d\mu_i \quad \textit{п.н. в} \quad t \in (\widetilde{t}_0, \widetilde{t}_1], \\ \mathcal{H}(\widetilde{t}_0, \widetilde{x}(\widetilde{t}_0), p(\widetilde{t}_0), \lambda) = -b'_{0_t}(\widetilde{t}_0, \widetilde{x}(\widetilde{t}_0)) \cdot \ell_0.\end{aligned}$$

Мерките μ_i, $i = 1, 2, \dots, k$, са с носители съответно множествата

$$T_i := \{t \in [\widetilde{t}_0, \widetilde{t}_1] : g_i(t, \widetilde{x}(t)) = 0\}.$$

Накрая ще демонстрираме Теорема 11.1 върху един пример.

Пример 11.1. Разглеждаме екстремалната задача

$$\int_0^1 \big(\dot{x}(t)^2 + x(t)^2\big)\, dt \longrightarrow \inf,\ x(0) = 1,\ x(t) \geq 0.$$

Решение. Представяме задачата във вида

$$\left|\begin{array}{l} \displaystyle\int_0^1 \big(x^2(t) + u^2(t)\big)\, dt \longrightarrow \inf \\ \dot{x}(t) = u(t), \quad t \in [0,1], \\ x(0) = 1, \\ -x(t) \leq 0, \quad t \in [0,1]. \end{array}\right.$$

Прилагаме Теорема 11.1 с $f(t,x,u) := x^2 + u^2$, $\varphi(t,x,u) := u$, $U = \mathbb{R}$, $b_0(x) := x - 1$, $b_1(x) := 0$ и $g(t,x) := -x$, като $m = n = s_0 = s_1 = 1$. Функцията на Понтрягин има вида

$$H(t,x,u,p,\lambda) := p\,u - \lambda(x^2 + u^2).$$

Ако задачата има локално решение $(\widetilde{x}(t), \widetilde{u}(t))$, като $\widetilde{x} \in AC[0,1]$ и $\widetilde{u}(t)$ е ограничена измерима функция, то за него съществуват според посочената теорема $\lambda \geq 0$, $p \in BV[0,1]$, която е непрекъсната отляво, и регулярна борелова мярка μ върху $[0,1]$ с носител множеството

$$T := \{t \in [0,1] : \widetilde{x}(t) = 0\},$$

такива, че е в сила спрегнатото уравнение

$$p(t) = -2\lambda \int_t^1 \widetilde{x}(\tau)\, d\tau + \int_t^1 d\mu, \quad t \in (0,1]. \tag{11.19}$$

Освен това според второто твърдение на теоремата за почти всяко $t \in [0,1]$ функцията $H(t, \widetilde{x}(t), u, p(t), \lambda)$ достига своята най-голяма стойност по u в $\mathbb{R}$ в $\widetilde{u}(t)$.

Ако $\lambda > 0$, то H е парабола по u с връх обърнат нагоре. В такъв случай H достига най-голямата си стойност във върха си, т.е. за $u = p/(2\lambda)$ и $\widetilde{u}(t) = p(t)/(2\lambda)$ п.н. Тогава

$$p(t) = 2\lambda\,\dot{\widetilde{x}}(t) \quad \text{п.н. в} \quad [0,1].$$

Оттук и (11.19) получаваме

$$\dot{\widetilde{x}}(t) = -\int_t^1 \widetilde{x}(\tau)\, d\tau + m(t) \quad \text{п.н. в} \quad [0,1], \tag{11.20}$$

където сме положили

$$m(t) := \frac{1}{2\lambda} \int_t^1 d\mu.$$

Функцията $m(t)$ е монотонно намаляваща, неотрицателна и непрекъсната отляво. Също така, тя е константа върху всеки интервал, върху който $\widetilde{x}(t) > 0$.

След като $\widetilde{x} \in AC[0,1]$ и $\widetilde{x}(0) = 1$, то от (11.20) следва

$$\widetilde{x}(t) = 1 - \int_0^t \left(\int_v^1 \widetilde{x}(\tau)\, d\tau \right) dv + \int_0^t m(v)\, dv, \quad t \in [0,1]. \tag{11.21}$$

Ще анализираме това уравнение. Разглеждаме следните случаи:

1) $T = \emptyset$,

2) $T = \{t_*\}$,

3) съществуват две различни точки $t_1, t_2 \in T$ такива, че $T \cap (t_1, t_2) = \emptyset$,

4) T съдържа затворен интервал, който не се изражда в точка.

Като използваме непрекъснатостта на $\widetilde{x}(t)$, лесно установяваме, че ако T съдържа поне две точки и за някоя стойност на t между тях имаме $\widetilde{x}(t) > 0$, то е налице случай 3). Следователно се реализира поне една от изброените горе възможности.

1) Ако T е празно, то $\widetilde{x}(t) > 0$ навсякъде в $[0,1]$ и $\mu = 0$. Тогава и $m(t) = 0$ в $[0,1]$ и (11.21) се свежда до

$$\widetilde{x}(t) = 1 - \int_0^t \left(\int_v^1 \widetilde{x}(\tau)\, d\tau \right) dv, \quad t \in [0,1]. \tag{11.22}$$

Следователно $\widetilde{x} \in C^2[0,1]$ и

$$\ddot{\widetilde{x}}(t) = \widetilde{x}(t), \quad t \in [0,1]. \tag{11.23}$$

Следователно

$$\widetilde{x}(t) = c_1 e^t + c_2 e^{-t}, \quad t \in [0,1], \tag{11.24}$$

с константи c_1 и c_2. Те се определят от две гранични условия. Едното е даденото по условие $\widetilde{x}(0) = 1$, а другото следва от (11.22) и е $\dot{\widetilde{x}}(1) = 0$. Така получаваме, че едно възможно решение на задачата е

$$\widetilde{x}(t) = \frac{1}{1+e^2}\, e^t + \frac{e^2}{1+e^2}\, e^{-t}.$$

2) Нека $\widetilde{x}(t) = 0$ само при $t = t_*$. Ако $t_* < 1$, то $m(t) = 0$ при $t \in (t_*, 1]$. Сега от (11.20) следва, че $\widetilde{x}(t)$ е строго намаляваща в $[t_*, 1]$ и след като

$\widetilde{x}(t_*) = 0$, то бихме получили $\widetilde{x}(t) < 0$ за $t > t_*$, което е невъзможно. Ако пък $t_* = 1$, то $m(t) = m_0 > 0$ при $t \in [0,1]$. От (11.21) следва, че

$$\widetilde{x}(t) = 1 - \int_0^t \left(\int_v^1 \widetilde{x}(\tau)\, d\tau \right) dv + m_0 t, \quad t \in [0,1].$$

Ако съществува $\widetilde{x}$, което удовлетворява това уравнение, то е от вида (11.24), като освен това удовлетворява условията $\widetilde{x}(0) = 1$, $\widetilde{x}(1) = 0$ и $\dot{\widetilde{x}}(1) = m_0$. Първите две гранични условия влекат

$$\widetilde{x}(t) = -\frac{1}{e^2 - 1}\, e^t + \frac{e^2}{e^2 - 1}\, e^{-t}.$$

Но тогава $m_0 = \dot{\widetilde{x}}(1) = -2e/(e^2 - 1) < 0$ — противоречие.

3) Ако за някои две различни точки t_1 и t_2, $0 < t_1 < t_2 \leq 1$, имаме $\widetilde{x}(t_1) = \widetilde{x}(t_2) = 0$ и $\widetilde{x}(t) > 0$ при $t \in (t_1, t_2)$, то $m(t) = \text{const}$ върху (t_1, t_2). Сега от (11.20) следва, че $\dot{\widetilde{x}}$ съвпада п.н. в $[t_1, t_2]$ с монотонно растяща функция. Следователно $\widetilde{x}(t)$ е изпъкнала върху $[t_1, t_2]$ и сега от $\widetilde{x}(t_1) = \widetilde{x}(t_2) = 0$ установяваме, че $\widetilde{x}(t) \leq 0$ за $t \in (t_1, t_2)$ — противоречие.

4) Накрая да разгледаме случая, в който за някои две различни точки t_1 и t_2, $0 < t_1 < t_2 \leq 1$, имаме $[t_1, t_2] \subseteq T$. Тогава $\widetilde{x}(t) = 0$ за $t \in [t_1, t_2]$. Следователно в интервала $[t_1, t_2]$ ще имаме

$$\begin{aligned}
&\dot{\widetilde{x}}(t) = 0, \\
&\int_t^1 \widetilde{x}(\tau)\, d\tau = \text{const}, \\
&m(t) \text{ строго намалява.}
\end{aligned}$$

Това е в противоречие с (11.20).

С това разгледахме случая $\lambda > 0$.

Ако $\lambda = 0$, то $H = pu$ и H е ограничена отгоре по u в $\mathbb{R}$, само ако $p(t) = 0$ п.н. в $[0,1]$. От (11.19) (с $\lambda = 0$) се вижда, че $p(t)$ е монотонно намаляваща в $(0,1]$. Следователно $p(t) = 0$ за всяко $t \in (0,1]$. Пак от (11.19) следва $\mu([t,1]) = 0$ за всяко $t \in (0,1]$, откъдето на свой ред следва, че $T \subseteq \{0\}$. След като $\widetilde{x}(0) \neq 0$, заключаваме, че $T = \emptyset$ и $\mu = 0$. Следователно $\widetilde{x}(t) > 0$ навсякъде в $[0,1]$. За останалите два множителя на Лагранж ℓ_0 и ℓ_1 теоремата не дава полезни зависимости. В този случай това необходимо условие не ни помага да намерим възможни решения на задачата, но ако за предполагаемото решение на задачата $\widetilde{x}(t)$ имаме $\widetilde{x}(t) > 0$ за всяко $t \in [0,1]$, то фазовото ограничение не играе роля (съществува околност на $\widetilde{x}$ в $C[0,1]$, в която стойността на функционала остава най-малка в $\widetilde{x}$). В такъв случай е приложим Принципът за максимума за задачи без фазови ограничения, формулиран в Теорема 9.3. Прилагаме тази теорема с $f(t,x,u) := x^2 + u^2$, $\varphi(t,x,u) := u$, $U = \mathbb{R}$,

$b_0(t,x) := (t, x-1)$ и $b_1(t,x) := t-1$, като $m = n = s_1 = 1$ и $s_0 = 2$. Функцията на Понтрягин има същия вид

$$H(t,x,u,p,\lambda) := p\,u - \lambda(x^2+u^2).$$

Ако задачата има локално решение $(\widetilde{x}(t), \widetilde{u}(t))$, като $\widetilde{x} \in AC[0,1]$ и $\widetilde{u}$ е ограничена измерима функция, дефинирана върху $[0,1]$, то за него съществуват според посочената теорема $\lambda \geq 0$, $\ell_0 := (\alpha, \beta) \in \mathbb{R}^2$, $\ell_1 \in \mathbb{R}$ и $p \in AC[0,1]$, не всичките нула, такива, че

$$\dot{p}(t) = 2\lambda\,\widetilde{x}(t) \quad \text{п.н. в} \quad [0,1], \tag{11.25}$$

$$p(0) = \beta, \tag{11.26}$$

$$p(1) = 0. \tag{11.27}$$

Също така според второто твърдение на теоремата за почти всяко $t \in [0,1]$ функцията $H(t, \widetilde{x}(t), u, p(t), \lambda)$ достига своята най-голяма стойност по u в $\mathbb{R}$ в $\widetilde{u}(t)$.

Ако допуснем, че $\lambda = 0$, получаваме от (11.25) и (11.27), че $p(t) \equiv 0$. Предвид (11.26) получаваме още $\beta = 0$. Също така щом $p(t) \equiv 0$ и $\lambda = 0$, то $H(t,x,u,p,\lambda) \equiv 0$, откъдето $\mathcal{H}(t,x,p,\lambda) \equiv 0$. Според Теорема 9.3(iii)

$$\mathcal{H}(0, \widetilde{x}(0), p(0), \lambda) = -\alpha, \quad \mathcal{H}(1, \widetilde{x}(1), p(1), \lambda) = \ell_1.$$

Следователно $\alpha = \ell_1 = 0$. Така получихме, че всички множители на Лагранж са нула, което е противоречие.

Ако $\lambda > 0$, то разсъжденията са на практика индентични с тези в случай 1) по-горе. Отново благодарение на Теорема 9.3(ii) установяваме, че $\widetilde{u}(t) = p(t)/(2\lambda)$ п.н. и следователно

$$p(t) = 2\lambda\,\dot{\widetilde{x}}(t) \quad \text{п.н. в} \quad [0,1]. \tag{11.28}$$

Тъй като $p, \widetilde{x} \in AC[0,1]$, основната теорема на диференциалното и интегралното смятане (теоремата на Лайбниц-Нютон) за лебеговия интеграл и (11.25) влекат $p \in C^1[0,1]$, както и че (11.25) е изпълнено навсякъде. След това чрез същия аргумент с разменени роли на p и $\widetilde{x}$ получаваме от (11.28), че $\widetilde{x} \in C^2[0,1]$, както и че (11.28) е също изпълнено навсякъде. Накрая (11.25) и (11.28) (изпълнени навсякъде) влекат (11.23). Граничните условия са даденото по условие $\widetilde{x}(0) = 1$ и още $\dot{\widetilde{x}}(1) = p(1) = 0$, което следва от (11.25) (изпълнено навсякъде) и (11.27). Така и в този случай достигнахме до същото $\widetilde{x}$, намерено в случай 1).

С това приключихме с разглеждането на случая $\lambda = 0$ за първоначално формулираната задача в примера.

Установихме, че единственото възможно решение на задачата е

$$\widetilde{x}(t) = \frac{1}{1+e^2}\,e^t + \frac{e^2}{1+e^2}\,e^{-t}.$$

Ще докажем, че то действително реализира и то глобален минимум на функционала в задачата. Нека $h \in AC[0,1]$ и $h(0) = 0$. Имаме

$$\int_0^1 \Big[(\dot{\widetilde{x}}(t) + \dot{h}(t))^2 + (\widetilde{x}(t) + h(t))^2\Big]\,dt = \int_0^1 \Big(\dot{\widetilde{x}}(t)^2 + \widetilde{x}(t)^2\Big)\,dt$$
$$+ 2\int_0^1 \dot{\widetilde{x}}(t)\,\dot{h}(t)\,dt + 2\int_0^1 \widetilde{x}(t)\,h(t)\,dt + \int_0^1 \Big(\dot{h}(t)^2 + h(t)^2\Big)\,dt.$$

Последният интеграл е неотрицателен. За сумата на втория и третия, след интегриране по части в първия, намираме

$$\int_0^1 \dot{\widetilde{x}}(t)\,\dot{h}(t)\,dt + \int_0^1 \widetilde{x}(t)\,h(t)\,dt$$
$$= \dot{\widetilde{x}}(1)h(1) - \dot{\widetilde{x}}(0)h(0) + \int_0^1 \Big(\widetilde{x}(t) - \ddot{\widetilde{x}}(t)\Big)\,h(t)\,dt = 0,$$

където сме използвали, че $\dot{\widetilde{x}}(1) = 0$, $h(0) = 0$ и (11.23).

Така установихме, че функцията

$$\widetilde{x}(t) = \frac{1}{1+e^2}\,e^t + \frac{e^2}{1+e^2}\,e^{-t}$$

е глобално решение на задачата.

В заключение да отбележим, че понеже $\widetilde{x}(t)$ не реализира за никое t фазовото ограничение като равенство, то можем да решим тази задача с помощта на Теорема 9.3. Дори т.к. решението ѝ принадлежи на $C^1[0,1]$, бихме го намерили посредством необходимото условие на Ойлер за задачата със свободен край (вж. Задача 1 в § 3) или дори чрез необходимото условие на Ойлер за задачата със закрепени краища. Преимуществото на Теореми 9.3 и 11.1 е в това, че ни дават възможност да намерим решенията в по-широкото пространство на абсолютно непрекъснатите функции с ограничена п.н. производна. А техният недостатък се състои в това, че понякога не ни помагат да намерим всички решения на екстремалната задача. Това се случва, когато не можем да изключим случая $\lambda = 0$. Това е особено слабост на Теорема 11.1.

Задачи

Решете екстремалните задачи:

1. $\displaystyle\int_0^1 \Big(\dot{x}^2(t) - x^2(t)\Big)\,dt \longrightarrow \inf,\ x(0) = x(1) = 0,\ |x(t)| \le 1;$

2. $\displaystyle\int_0^1 x(t)\,dt \longrightarrow \sup,\ x(0) = x(1) = 0,\ x(t) \le 1,\ |\dot{x}(t)| \le 1.$

12 Достатъчни условия за екстремум

Нека X е топологично линейно пространство, $C \subseteq X$ и $f : X \to \mathbb{R}$. Разглеждаме екстремалната задача

$$\left| \begin{array}{l} f(x) \longrightarrow \inf \\ x \in C. \end{array} \right. \tag{12.1}$$

Ще докажем едно изключително елементарно, но полезно достатъчно условие за нейните решения.

Дефиниция 12.1. Казваме, че функцията $\varphi : X \to \mathbb{R}$ е *K-функция* за задачата (12.1) в точката $\widetilde{x} \in X$, ако:

(а) $f(\widetilde{x}) = \varphi(\widetilde{x})$;

(б) $\varphi(\widetilde{x}) \le \varphi(x), \quad x \in C$;

(в) $\varphi(x) \le f(x), \quad x \in C$.

Казваме, че функцията $\varphi : X \to \mathbb{R}$ е *локална K-функция* за задачата (12.1) в точката $\widetilde{x} \in X$, ако свойства (б) и (в) са изпълнени в околност на $\widetilde{x}$.

Следва едно елементарно достатъчно условие за решение на разглежданата екстремална задача, формулирано в термините на K-функция.

Твърдение 12.2. *Ако съществува (локална) K-функция за задачата* (12.1) *в точката $\widetilde{x} \in X$, то $\widetilde{x}$ е (локално) решение на* (12.1).

Доказателство. Нека $\varphi(x)$ е (локална) K-функция на задачата (12.1) в точката $\widetilde{x} \in X$. Непосредствено от дефиницията на K-функция получаваме за $x \in C$, съответно $x \in C \cap U$, където $U \subset X$ е околност на $\widetilde{x}$:

$$f(\widetilde{x}) \overset{(а)}{=} \varphi(\widetilde{x}) \overset{(б)}{\le} \varphi(x) \overset{(в)}{\le} f(x).$$

□

Бележка 12.3. Нека $\varphi(x)$ е K-функция на задачата (12.1) в точката $\widetilde{x} \in X$ (**глобална**). Тогава, както установихме току-що, $\widetilde{x}$ е **глобално** решение на (12.1). Нека $\overline{x} \in X$ е друго **глобално** решение на задачата. Тогава от дефиницията на K-функция следва, че

$$f(\overline{x}) = f(\widetilde{x}) \overset{(а)}{=} \varphi(\widetilde{x}) \overset{(б)}{\le} \varphi(\overline{x}) \overset{(в)}{\le} f(\overline{x}).$$

Следователно

$$\varphi(\widetilde{x}) = \varphi(\overline{x}) = f(\overline{x}),$$

откъдето получаваме, че функцията $\varphi(x)$ притежава свойствата (а)-(в) с $\overline{x}$ на мястото на $\widetilde{x}$. Следователно $\varphi(x)$ е K-функция за задачата (12.1) в точката $\overline{x} \in X$. Така установихме, че ако дадена функция е K-функция за задачата (12.1) в някое нейно решение, то тя е K-функция във всяко друго решение. Това ни дава основания да говорим за *K-функция за задачата*. Да напомним, че проведените тук разсъждения касаят само глобалните решения.

Достатъчното условие за минимум в Твърдение 12.2 се установява лесно, но не представлява ефективен начин за решаване конкретни екстремални задачи. Неговото значение по-скоро се състои в това, че ни предоставя общ метод за установяване на достатъчни условия за екстремум. Тук ще разгледаме две такива негови приложения: за изпъкнали и за гладки задачи.

12.1 Изпъкнали задачи

Нека X е топологично линейно пространство, $f_i : X \to \mathbb{R}$, $i = 0, 1, \dots, n$, са изпъкнали функции и $A \subseteq X$ е изпъкнало множество. Разглеждаме задачата

$$
\left|\begin{array}{ll}
f_0(x) \longrightarrow \inf & \\
f_i(x) \le 0, \quad i = 1, 2, \dots, n, & (*) \\
x \in A. & (**)
\end{array}\right. \tag{12.2}
$$

За $\widetilde{x} \in X$ дефинираме функцията

$$
\varphi(x) := f_0(\widetilde{x}) - \sum_{i=1}^{n} \lambda_i f_i(x),
$$

където

$$
\lambda_i \ge 0, \quad \lambda_i f_i(\widetilde{x}) = 0, \quad i = 1, 2, \dots, n. \tag{12.3}
$$

Да положим $C := \{x \in A : f_i(x) \le 0,\ i = 1, 2, \dots, n\}$. Очевидно $\varphi(x)$ удовлетворява свойства (а) и (б) на дефиницията за K-функция. Остава да се покаже, че за някое $\widetilde{x}$ множителите λ_i, удовлетворяващи (12.3), може така да се определят, че да имаме

$$
\varphi(x) \le f_0(x), \quad x \in A.
$$

Разбира се, за да постигнем това по-лесно, е полезно да знаем най-малката стойност на $f_0(x)$ при условията $(*)$-$(**)$ и кои от условията $(*)$ се изпълняват като равенства в точка на минимум.

Образуваните по горния начин K-функции за изпъкналата задача се наричат *стандартни*.

12.2 Гладки задачи

Преминаваме към едно следствие на Твърдение 12.2 за гладки задачи. В неговата формулировка основна роля играе втората производна по Фреше.

Теорема 12.4. *Нека $f : X \to \mathbb{R}$ и $F : X \to Y$, където X и Y са банахови пространства. Разглеждаме екстремалната задача*

$$f(x) \longrightarrow \inf, \quad F(x) = 0. \tag{12.4}$$

За $x \in X$ и $y^* \in Y^*$ полагаме

$$\mathcal{L}(x, y^*) := f(x) + \langle y^*, F(x) \rangle.$$

Нека точката $\widetilde{x} \in X$ удовлетворява следните условия:

(а) изображенията $f(x)$ и $F(x)$ притежават първа и втора производна на Фреше в околност на $\widetilde{x}$, като вторите производни са непрекъснати в $\widetilde{x}$;

(б) $F(\widetilde{x}) = 0$;

(в) $\operatorname{Im} F'(\widetilde{x}) = Y$;

(г) съществуват $\widetilde{y}^* \in Y^*$ и $\alpha > 0$ такива, че:

$$\mathcal{L}'_x(\widetilde{x}, \widetilde{y}^*) = 0$$

и

$$\mathcal{L}''_{xx}(\widetilde{x}, \widetilde{y}^*)(\xi, \xi) \geq \alpha \, \|\xi\|_X^2 \quad \forall \xi \in \operatorname{Ker} F'(\widetilde{x}).$$

Тогава $\widetilde{x}$ е локално решение на задачата (12.4).

Бележка 12.5. Да припомним, че $\mathcal{L}(x, y^*)$ е функцията на Лагранж за екстремалната задача (12.4) и според Теорема 4.6 условието да съществува $\widetilde{y}^* \in Y^*$ такова, че $\mathcal{L}'_x(\widetilde{x}, \widetilde{y}^*) = 0$ е необходимо, за да бъде точката $\widetilde{x}$ нейно решение. Може да се покаже, че ако са налице всички направени предположения в Теорема 12.4 без непременно второто в (г) и точката $\widetilde{x}$ е решение на задача (12.4), то $\mathcal{L}''_{xx}(\widetilde{x}, \widetilde{y}^*)(\xi, \xi) \geq 0$ за всяко $\xi \in \operatorname{Ker} F'(x)$.

В доказателството на теоремата ще използваме следното помощно твърдение.

Лема 12.6. *Нека X и Y са банахови пространства, $F : X \to Y$ е диференцируемо по Фреше в точката $x_0 \in X$, като $\operatorname{Im} F'(x_0) = Y$, и нека $F(x_0) = 0$. Тогава за всяко $\varepsilon > 0$ съществува $\delta > 0$ такова, че ако $\|\zeta - x_0\|_X < \delta$ и $F(\zeta) = 0$, то съществува $\eta \in X$ такова, че $\zeta - x_0 - \eta \in \operatorname{Ker} F'(x_0)$ и $\|\eta\|_X \leq \varepsilon \|\zeta - x_0\|_X$.*

Доказателство. Твърдението е тривиално за $\zeta = x_0$ — можем да вземем $\eta = 0$. По-нататък, ще считаме, че $\zeta \neq x_0$.

Нека с $B_Z(z_0, r)$ означим отвореното кълбо в нормираното линейно пространство Z с център точката z_0 и радиус $r > 0$. От теоремата за отвореното изображение[1] следва, че съществува $\delta_1 > 0$ такова, че

$$F'(x_0)\big(B_X(0,1)\big) \supseteq B_Y(0, \delta_1). \tag{12.5}$$

Нека $\varepsilon > 0$ е произволно фиксирано. Щом $F(x)$ е диференцируемо по Фреше в x_0, то съществува $\delta > 0$ такова, че ако $\|\zeta - x_0\|_X < \delta$, то

$$\|F(\zeta) - F(x_0) - F'(x_0)(\zeta - x_0)\|_Y < \varepsilon\, \delta_1 \|\zeta - x_0\|_X.$$

Ако $F(\zeta) = 0$ и щом $F(x_0) = 0$, то горното неравенство придобива вида

$$\|F'(x_0)(\zeta - x_0)\|_Y < \varepsilon\, \delta_1 \|\zeta - x_0\|_X.$$

След като разделим двете страни на $\varepsilon\, \|\zeta - x_0\|_X$, получаваме

$$\Big\|F'(x_0)\big[\big(\varepsilon\, \|\zeta - x_0\|_X\big)^{-1}(\zeta - x_0)\big]\Big\|_Y < \delta_1.$$

Сега, като вземем предвид (12.5), заключаваме, че съществува $\widetilde{\eta} \in X$ с $\|\widetilde{\eta}\|_X < 1$ такова, че

$$F'(x_0)\widetilde{\eta} = F'(x_0)\big[\big(\varepsilon\, \|\zeta - x_0\|_X\big)^{-1}(\zeta - x_0)\big]. \tag{12.6}$$

Полагаме $\eta := \varepsilon\, \|\zeta - x_0\|_X \widetilde{\eta}$. Тогава $\|\eta\|_X < \varepsilon \|\zeta - x_0\|_X$ и благодарение на (12.6) имаме $F'(x_0)\big(\eta - (\zeta - x_0)\big) = 0$, което означава, че $\zeta - x_0 - \eta \in \operatorname{Ker} F'(x_0)$. □

Готови сме да преминем към доказателството на Теорема 12.4.

Доказателство на Теорема 12.4. Да означим с U околността на точката $\widetilde{x}$, в която са изпълнени направените в (а) предположения. Прилагаме теоремата на Люстерник към изображението $F(x)$ в точката $x_0 = \widetilde{x}$. Така получаваме, че съществува околност $V \subseteq U$ на $\widetilde{x}$, константа $c > 0$ и изображение $h : V \to X$ такива, че

$$F(x + h(x)) = 0, \tag{12.7}$$

$$\|h(x)\|_X \leq c \|F(x)\|_Y \tag{12.8}$$

за всяко $x \in V$. Да припомоним, че според (б) $F(\widetilde{x}) = 0$. Оттук нататък ще предполагаме винаги, че $x \in V$.

[1] Теоремата за отвореното изображение гласи, че ако X и Y са баханови пространства и $\Lambda : X \to Y$ е ограничен линеен оператор, като $\Lambda(X) = Y$, то образът на всяко отворено множество в X, под действието на Λ, е отворено множество в Y.

Ще покажем, че функцията

$$\varphi(x) := f(\widetilde{x}) - \langle \widetilde{y}^*, F(x) \rangle - \|h(x)\|_X$$

е локална K-функция за задачата (12.4) в точката $\widetilde{x}$. Оттук, благодарение на Твърдение 12.2 с $C = \operatorname{Ker} F$, веднага следва, че $\widetilde{x}$ е локално решение на екстремалната задача.

Да забележим, че ако $F(x) = 0$, то от (12.8) следва, че $h(x) = 0$ и тогава $\varphi(x) = f(\widetilde{x})$. Следователно, функцията $\varphi(x)$ притежава свойства (а) и (б) в Дефиниция 12.1 при $x \in C \cap V$. Остава да проверим свойство (в), т.е. да покажем, че съществува околност V' на $\widetilde{x}$ такава, че $\varphi(x) \leq f(x)$ при $x \in C \cap V'$.

Да положим

$$g(x) := \mathcal{L}(x, \widetilde{y}^*) - f(\widetilde{x}).$$

Тогава

$$f(x) - \varphi(x) = g(x) + \|h(x)\|_X. \tag{12.9}$$

Според направените в (а), (б) и (г) предположения, функцията $g(x)$ притежава първа и втора производна на Фреше във V, при това

$$g(\widetilde{x}) = 0, \quad g'(\widetilde{x}) = 0, \quad g''(\widetilde{x}) = \mathcal{L}''_{xx}(\widetilde{x}, \widetilde{y}^*)$$

и следователно

$$g(x) = \frac{1}{2}\,\mathcal{L}''_{xx}(\widetilde{x}, \widetilde{y}^*)(x - \widetilde{x}, x - \widetilde{x}) + o(\|x - \widetilde{x}\|_X^2). \tag{12.10}$$

Представяме разликата на $f(x)$ и $\varphi(x)$ във вида (вж. (12.9))

$$f(x) - \varphi(x) = [g(x) - g(x + h(x))] + \|h(x)\|_X + g(x + h(x)). \tag{12.11}$$

Според теоремата за крайните нараствания имаме

$$|g(x) - g(x + h(x))| \leq \sup_{t \in [0,1]} \|g'(x + th(x))\| \, \|h(x)\|_X. \tag{12.12}$$

Ще покажем, че можем да направим $\|g'(x + th(x))\|$ колкото желаем малка равномерно по $t \in [0, 1]$, стига да вземем x достатъчно близо до $\widetilde{x}$. За всяко $t \in [0, 1]$, като използваме (12.8), установяваме, че

$$\begin{aligned} \|x + th(x) - \widetilde{x}\|_X &\leq \|x - \widetilde{x}\|_X + \|h(x)\|_X \\ &\leq \|x - \widetilde{x}\|_X + c\,\|F(x)\|_Y. \end{aligned}$$

Сега, като вземем още предвид, че $F(x)$ е непрекъснато в $\widetilde{x}$ и $F(\widetilde{x}) = 0$, заключаваме, че можем да направим $x + th(x)$ колкото желаем близо до

$\widetilde{x}$ и то равномерно по $t \in [0,1]$, стига да вземем x достатъчно близо до $\widetilde{x}$; казано по друг начин:

$$\forall \delta > 0 \; \exists \delta' > 0 : \|x - \widetilde{x}\|_X < \delta' \implies \|x + th(x) - \widetilde{x}\|_X < \delta \quad \forall t \in [0,1]. \tag{12.13}$$

От друга страна производната $g'(x)$ е непрекъсната в операторната норма, както следва от (а), и $g'(\widetilde{x}) = 0$. Като вземем още предвид (12.13), установяваме, че за всяко $\varepsilon > 0$ съществува околност V_1 на $\widetilde{x}$ такава, че

$$\|g'(x + th(x))\| < \varepsilon \quad \forall t \in [0,1] \quad \forall x \in V_1.$$

Сега, като се върнем към (12.12), заключаваме, че

$$|g(x) - g(x + h(x))| \le \varepsilon \, \|h(x)\|_X \quad \forall x \in V_1. \tag{12.14}$$

Ще оценим последното събираемо в дяснана страна на (12.11). За краткост полагаме $\zeta := \zeta(x) := x + h(x)$. От формула (12.10) с ζ вместо x получаваме

$$g(\zeta) = \frac{1}{2}\, \mathcal{L}''_{xx}(\widetilde{x}, \widetilde{y}^*)(\zeta - \widetilde{x}, \zeta - \widetilde{x}) + o(\|\zeta - \widetilde{x}\|_X^2). \tag{12.15}$$

Според (12.7) имаме, че, ако $x \in V$, то $F(\zeta) = 0$. Прилагаме Лема 12.6 с $x_0 = \widetilde{x}$. Така установяваме, че за всяко $\varepsilon > 0$ съществува околност V_2 на $\widetilde{x}$ такава, че ако $\zeta \in V_2$, то съществува $\eta \in X$, зависещо от x, такова, че

$$\zeta - \widetilde{x} - \eta \in \operatorname{Ker} F'(\widetilde{x}) \quad \text{и} \quad \|\eta\|_X < \varepsilon \|\zeta - \widetilde{x}\|_X. \tag{12.16}$$

От (12.13) с $t = 1$ следва, че съществува околност V_3 на $\widetilde{x}$ такава, че ако $x \in V_3$, то $\zeta \in V_2$. Тогава имаме, че за всяко фиксирано $x \in V_3$ съществува $\eta \in X$, удовлятворяващо (12.16).

Като използваме (г), както и ограничеността, билинейността и симетричността на втората производна, получаваме с $d := \|\mathcal{L}''_{xx}(\widetilde{x}, \widetilde{y}^*)\|$

$$\begin{aligned}
\mathcal{L}''_{xx}(\widetilde{x}, \widetilde{y}^*)(\zeta - \widetilde{x}, \zeta - \widetilde{x}) &= \mathcal{L}''_{xx}(\widetilde{x}, \widetilde{y}^*)(\zeta - \widetilde{x} - \eta, \zeta - \widetilde{x} - \eta) \\
&\quad + 2\mathcal{L}''_{xx}(\widetilde{x}, \widetilde{y}^*)(\eta, \zeta - \widetilde{x} - \eta) + \mathcal{L}''_{xx}(\widetilde{x}, \widetilde{y}^*)(\eta, \eta) \\
&\ge \alpha \, \|\zeta - \widetilde{x} - \eta\|_X^2 - 2d \, \|\eta\|_X \, \|\zeta - \widetilde{x} - \eta\|_X - d \, \|\eta\|_X^2 \\
&\ge \alpha \big(\|\zeta - \widetilde{x}\|_X - \|\eta\|_X\big)^2 - 2d \, \|\eta\|_X \big(\|\zeta - \widetilde{x}\|_X + \|\eta\|_X\big) - d \, \|\eta\|_X^2.
\end{aligned}$$

Сега за произволно фиксирано $x \in V_3$ и съответващото му η получаваме от последното неравенство и неравенството в (12.16), оценката

$$\mathcal{L}''_{xx}(\widetilde{x}, \widetilde{y}^*)(\zeta - \widetilde{x}, \zeta - \widetilde{x}) \ge \big[\alpha(1 - \varepsilon)^2 - 2d\varepsilon(1 + \varepsilon) - d\varepsilon^2\big] \, \|\zeta - \widetilde{x}\|_X^2.$$

Оттук и (12.15), като използваме още (12.13), установяваме, че съществува околност V_4 на $\widetilde{x}$ такава, че ако $x \in V_4$, то

$$g(x+h(x)) \geq \frac{1}{2}\left[\alpha(1-\varepsilon)^2 - 2d\varepsilon(1+\varepsilon) - d\varepsilon^2 - \varepsilon\right] \|x+h(x)-\widetilde{x}\|_X^2. \tag{12.17}$$

Накрая, като комбинираме (12.11), (12.14) и (12.17), установяваме, че за всяко $\varepsilon > 0$ съществува околност V' на $\widetilde{x}$ такава, че за $x \in V'$ имаме

$$\begin{aligned} f(x) - \varphi(x) \geq (1-\varepsilon)\|h(x)\|_X& \\ + \frac{1}{2}\left[\alpha(1-\varepsilon)^2 - 2d\varepsilon(1+\varepsilon) - d\varepsilon^2 - \varepsilon\right] &\|x + h(x) - \widetilde{x}\|_X^2. \end{aligned}$$

За да приключим доказателството, остава да фиксираме $\varepsilon \in (0,1)$ така, че изразът в квадратните скоби да е неотрицателен (а това е така за достатъчно малки $\varepsilon > 0$). $\square$

Пример 12.1. Ще демонстрираме Теорема 12.4 в намирането на локални решения на екстремалната задача

$$\int_0^1 \dot{x}(t)^2\, dt \longrightarrow \inf, \quad x(0) = 0,\ x(1) = 1. \tag{12.18}$$

Ще потърсим решението ѝ в естествената дефиниционна област на функционала, а именно множеството от абсолютно непрекъснатите функции върху интервала $[0,1]$, квадратите на чиито първи производни са интегруеми по Лебег върху $[0,1]$. Това множество от функции се означава чрез $W_2^1[0,1]$. Снабдено с нормата

$$\|x\|_{W_2^1[0,1]} := \|x\|_{L_2[0,1]} + \|\dot{x}\|_{L_2[0,1]} = \left(\int_0^1 x(t)^2\, dt\right)^{1/2} + \left(\int_0^1 \dot{x}(t)^2\, dt\right)^{1/2},$$

то е банахово пространство.

Прилагаме теоремата с $X := W_2^1[0,1]$, $Y := \mathbb{R}^2$, $f(x) := \displaystyle\int_0^1 \dot{x}(t)^2\, dt$ и $F(x) := (x(0), x(1) - 1)$.

За да установим (a) по отношение на $f(x)$, използваме, че за $x, h \in W_2^1[0,1]$ имаме

$$f(x+h) = f(x) + 2\int_0^1 \dot{x}(t)\dot{h}(t)\, dt + \int_0^1 \dot{h}(t)^2\, dt.$$

Забелязваме, че

$$\int_0^1 \dot{h}(t)^2\, dt = \|\dot{h}\|_{L_2[0,1]}^2 = o(\|h\|_{W_2^1[0,1]}).$$

За всяко фиксирано $x \in W_2^1[0,1]$ интегралът

$$\int_0^1 \dot{x}(t)\dot{h}(t)\,dt$$

дефинира линеен ограничен функционал на h в $W_2^1[0,1]$. Ограничеността се установява с помощта на неравенството на Коши.

Аналогично
$$\int_0^1 \dot{h}_1(t)\,\dot{h}_2(t)\,dt$$
дефинира билинеен симетричен ограничен функционал на променливата (h_1,h_2) в $W_2^1[0,1] \times W_2^1[0,1]$.

Следователно

$$f'(x)h = 2\int_0^1 \dot{x}(t)\dot{h}(t)\,dt, \quad f''(x)(h,h) = 2\int_0^1 \dot{h}(t)^2\,dt.$$

Функционалът $f''(x)$ не зависи от x. Следователно тривиално е непрекъснат по x в операторната норма.

Относно изображението $F(x)$ имаме за $x,h \in W_2^1[0,1]$

$$F(x+h) = F(x) + (h(0),h(1)).$$

Ще докажем, че изображението, което на $h \in W_2^1[0,1]$ съпоставя точката $(h(0),h(1)) \in \mathbb{R}^2$, е ограничено. Очевидно то е линейно. Нека считаме за определеност, че $\mathbb{R}^2$ е снабдено с евклидовата норма, но да напомним, че в крайномерно пространство всеки две норми са еквивалентни. С това ще сме показали, че за всяко $x \in W_2^1[0,1]$

$$F'(x)h = (h(0),h(1)), \quad F''(x) = 0.$$

За да докажем, че изображението, което на $h \in W_2^1[0,1]$ съпоставя $(h(0),h(1)) \in \mathbb{R}^2$, е ограничено, е достатъчно да се уверим, че съществува положителна константа c_0 такава, че

$$|h(0)|, |h(1)| \le c_0\|h\|_{W_2^1[0,1]}, \quad h \in W_2^1[0,1].$$

За тази цел, от своя страна, е достатъчно да докажем, че

$$|h(0)| \le \|h\|_{W_2^1[0,1]}, \quad h \in W_2^1[0,1], \tag{12.19}$$

защото, както непосредствено се вижда,

$$\begin{aligned} |h(1)-h(0)| &= \left|\int_0^1 \dot{h}(t)\,dt\right| \le \int_0^1 |\dot{h}(t)|\,dt \\ &\le \left(\int_0^1 \dot{h}(t)^2\,dt\right)^{1/2} = \|\dot{h}\|_{L_2[0,1]} \le \|h\|_{W_2^1[0,1]}, \end{aligned} \tag{12.20}$$

като при втората оценка използваме неравенството на Коши (или пък това на Йенсен).

Да докажем (12.19). Ако съществува $\xi \in [0,1]$ такова, че $h(\xi) = 0$, то аналогично на (12.20) получаваме

$$\begin{aligned}|h(0)| = \left|\int_0^{\xi} \dot{h}(t)\,dt\right| &\leq \int_0^{\xi} |\dot{h}(t)|\,dt \\ &\leq \int_0^1 |\dot{h}(t)|\,dt \leq \|h\|_{W_2^1[0,1]}.\end{aligned}$$

С това установихме (12.19) при предположението, че $h(t)$ има нула.

Ако $h(t)$ няма нула в $[0,1]$, то, понеже е непрекъсната функция, тя приема или само положителни стойности, или само отрицателни. Ще разгледаме първия случай — вторият се свежда към него, като вместо h вземем $-h$. Като използваме отново непрекъснатостта на $h(t)$, заключаваме, че $h(t)$ има най-малка стойност в $[0,1]$ и тя е положителна. Да означим някоя от стойностите на аргумента, в която тя се достига, с ξ. Тогава $h(t) \geq h(\xi) > 0$ за всяко $t \in [0,1]$. Аналогично на (12.20), показваме, че

$$h(0) - h(\xi) \leq \|\dot{h}\|_{L_2[0,1]},$$

откъдето получаваме

$$h(0) \leq h(\xi) + \|\dot{h}\|_{L_2[0,1]}. \tag{12.21}$$

Остава да забележим, че

$$\|h\|_{L_2[0,1]} = \left(\int_0^1 h(t)^2\,dt\right)^{1/2} \geq h(\xi),$$

за да извлечем от (12.21) неравенството (12.19) в този случай.

С това проверихме, че са налице условията в (а) на Теорема 12.4 във всяка точка $\widetilde{x}$. Условие (в) е също тривиално изпълнено във всяка точка $\widetilde{x}$.

Функцията на Лагранж за разглежданата екстремална задача се дефинира чрез

$$\mathcal{L}(x, \ell_1, \ell_2) := f(x) + \ell_1 x(0) + \ell_2 (x(1) - 1),$$

където $\ell_1, \ell_2 \in \mathbb{R}$. Тук използваме, че ограничените линейни функционали y^* над $Y = \mathbb{R}^2$ се представят във вида

$$\langle y^*, y\rangle = \ell_1 y_1 + \ell_2 y_2, \quad y = (y_1, y_2) \in \mathbb{R}^2,$$

с някакви $\ell_1, \ell_2 \in \mathbb{R}$.

Както вече припомнихме (вж. Бележка 12.5), условието за дадено $\widetilde{x} \in W_2^1[0,1]$, удовлетворяващо граничните условия, да съществуват $\widetilde{\ell}_1, \widetilde{\ell}_2 \in \mathbb{R}$ такива, че

$$\mathcal{L}'_x(\widetilde{x}, \widetilde{\ell}_1, \widetilde{\ell}_2)h = 0 \quad \forall h \in W_2^1[0,1] \tag{12.22}$$

е необходимо, за да бъде $\widetilde{x}$ решение на задачата. Нека намерим функциите $\widetilde{x}$ с това свойство. За $h \in W_2^1[0,1]$ имаме

$$\begin{aligned}
\mathcal{L}'_x(\widetilde{x}, \widetilde{\ell}_1, \widetilde{\ell}_2)h &= 2\int_0^1 \dot{\widetilde{x}}(t)\dot{h}(t)\,dt + \widetilde{\ell}_1 h(0) + \widetilde{\ell}_2 h(1) \\
&= 2\int_0^1 \dot{\widetilde{x}}(t)\dot{h}(t)\,dt + \widetilde{\ell}_1 h(0) + \widetilde{\ell}_2 \left(h(0) + \int_0^1 \dot{h}(t)\,dt\right) \\
&= \int_0^1 (2\,\dot{\widetilde{x}}(t) + \widetilde{\ell}_2)\dot{h}(t)\,dt + (\widetilde{\ell}_1 + \widetilde{\ell}_2)h(0).
\end{aligned}$$

Забелязваме, че за да е изпълнено (12.22), трябва да имаме

$$\int_0^1 (2\,\dot{\widetilde{x}}(t) + \widetilde{\ell}_2)\dot{h}(t)\,dt = 0 \quad \forall h \in W_2^1[0,1] \quad \text{и} \quad \widetilde{\ell}_1 + \widetilde{\ell}_2 = 0.$$

Действително, ако допуснем, че $\widetilde{\ell}_1 + \widetilde{\ell}_2 \neq 0$, то с $h(t) := 1$, веднага получаваме противоречие с (12.22).

По-нататък забелязваме, че условието

$$\int_0^1 (2\,\dot{\widetilde{x}}(t) + \widetilde{\ell}_2)\dot{h}(t)\,dt = 0 \quad \forall h \in W_2^1[0,1]$$

влече $2\,\dot{\widetilde{x}}(t) + \widetilde{\ell}_2 = 0$ п.н. Действително, ако допуснем, че $2\,\dot{\widetilde{x}}(t) + \widetilde{\ell}_2 > 0$ върху някакво подмножество S на $[0,1]$ с положителна мярка, то с функцията

$$h(t) := \int_0^t \chi_S(\tau)\,d\tau$$

имаме $h \in AC[0,1]$, защото $\chi_S \in L[0,1]$, $\dot{h} = \chi_S$ и следователно $h \in W_2^1[0,1]$ и още

$$\int_0^1 (2\,\dot{\widetilde{x}}(t) + \widetilde{\ell}_2)\dot{h}(t)\,dt = \int_0^1 (2\,\dot{\widetilde{x}}(t) + \widetilde{\ell}_2)\chi_S(t)\,dt = \int_S (2\,\dot{\widetilde{x}}(t) + \widetilde{\ell}_2)\,dt > 0,$$

което е в противоречие с (12.22). Следователно $2\,\dot{\widetilde{x}}(t) + \widetilde{\ell}_2 \leq 0$ п.н. Аналогично се установява, че $2\,\dot{\widetilde{x}}(t) + \ell_2 \geq 0$ п.н.

Така доказахме, че $2\,\dot{\widetilde{x}}(t) + \widetilde{\ell}_2 = 0$ п.н. Това означава, че $\dot{\widetilde{x}}(t) = -\widetilde{\ell}_2/2$ п.н., откъдето предвид на това, че $\widetilde{x} \in AC[0,1]$, следва, че $\widetilde{x}(t)$ е линейна функция. Тя трябва да удовлетворява граничните условия $\widetilde{x}(0) = 0$ и

$\widetilde{x}(1) = 1$ (т.е. условие (б) в теоремата). Оттук окончателно получаваме, че решението на екстремалната задача (ако изобщо има такова) може да бъде само функцията $\widetilde{x}(t) := t$, като съответните $\widetilde{\ell}_1$ и $\widetilde{\ell}_2$ са $\widetilde{\ell}_1 = 2$ и $\widetilde{\ell}_2 = -2$.

Остава да покажем, че $\widetilde{x}(t)$ действително е решение. За тази цел ще се уверим, че тя удовлетворява второто условие в (г). Тогава от теоремата ще следва, че $\widetilde{x}(t)$ е решение на (12.18).

За втората производна на $\mathcal{L}$ по x в $(\widetilde{x}, \widetilde{\ell}_1, \widetilde{\ell}_2)$ имаме

$$\mathcal{L}''_{xx}(\widetilde{x}, \widetilde{\ell}_1, \widetilde{\ell}_2)(h,h) = 2\int_0^1 \dot{h}(t)^2\,dt = 2\|\dot{h}\|^2_{L_2[0,1]}.$$

За да установим, че тя притежава свойството в (г), трябва да покажем, че съществува константа c_1 такава, че

$$\|h\|_{W_2^1[0,1]} \le c_1\|\dot{h}\|_{L_2[0,1]} \quad \forall h \in T,$$

където сме положили

$$T := \operatorname{Ker} F'(\widetilde{x}) = \{h \in W_2^1[0,1] : h(0) = h(1) = 0\}.$$

Това неравенство на свой ред следва от

$$\|h\|_{L_2[0,1]} \le c_2\|\dot{h}\|_{L_2[0,1]} \quad \forall h \in T \tag{12.23}$$

с някаква константа c_2. За да установим последното, взимаме предвид, че ако $h \in AC[0,1]$ и $h(0) = 0$, то

$$h(t) = \int_0^t \dot{h}(\tau)\,d\tau$$

и следователно

$$\begin{aligned}
\|h\|_{L_2[0,1]} &= \left(\int_0^1 h(t)^2 dt\right)^{1/2} = \left(\int_0^1 \left(\int_0^t \dot{h}(\tau)\,d\tau\right)^2 dt\right)^{1/2} \\
&\le \left(\int_0^1 \left(\int_0^t |\dot{h}(\tau)|\,d\tau\right)^2 dt\right)^{1/2} \le \left(\int_0^1 \left(\int_0^1 |\dot{h}(\tau)|\,d\tau\right)^2 dt\right)^{1/2} \\
&= \int_0^1 |\dot{h}(\tau)|\,d\tau \le \left(\int_0^1 \dot{h}(\tau)^2\,d\tau\right)^{1/2} = \|\dot{h}\|_{L_2[0,1]},
\end{aligned}$$

като при последната оценка използваме неравенството на Коши (или пък това на Йенсен).

С това (12.23) е установено и от Теорема 12.4 следва, че $\widetilde{x}(t)$ е локално решение на задачата (12.18).

Последното разсъждение ни подсказва един елементарен начин да покажем, че $\widetilde{x}(t)$ е решение на задачата, без да прибягваме до Теорема 12.4. При това с негова помощ се установява дори че тази функция е глобален минимум.

Както по-горе, посредством неравенството на Коши, за всяка функция $x \in W_2^1[0,1]$, удовлетворяваща граничните условия $x(0) = 0$, $x(1) = 1$, получаваме

$$1 = [x(1) - x(0)]^2 = \left(\int_0^1 \dot{x}(t)\,dt \right)^2 \leq \int_0^1 \dot{x}(t)^2\,dt.$$

От друга страна за $\widetilde{x}(t)$ имаме

$$\int_0^1 \dot{\widetilde{x}}(t)^2 dt = 1.$$

Теорема 4.6 показва, че други решения, включително и локални, няма.

Задачи

1. Изведете, като частен случай на Теорема 12.4, достатъчно условие за локален минимум/максимум за изображения, дефинирани в $\mathbb{R}^n$, със стойности в $\mathbb{R}^m$.

2. Изведете, като частен случай на Теорема 12.4, достатъчно условие за слаб локален минимум/максимум за задачата със закрепени краища.

3. Покажете, че функцията $\widetilde{x}(t) := t$ е слабо локално решение[1] на екстремалната задача от Пример 12.1, като използвате уравнението на Ойлер-Лагранж за задачите със закрепени краища. Не пропускайте да установете, че действително тази функция е решение на задачата. Вярно ли е, че това решение е силно?[2] Сравнете тези резултати с установеното в Пример 12.1

4. Докажете твърдението, изказано в края на Бележка 12.5.

[1]Слабо в смисъл, че околностите се разглеждат в нормата на $C^1[0,1]$.
[2]Околностите се разглеждат в нормата на $C[0,1]$.

13 Конкретни екстремални задачи

Тук ще решим няколко конкретни екстремални задачи, като приложим изучените методи.

Преди да приложим някоя обща теорема към екстремална задача, добре е първо да проверим дали тя няма елементарно решение. Ето един такъв случай (вж. Пример 9.1)

Задача 1. Решете екстремалната задача

$$\int_0^1 x(t)\,dt \longrightarrow \inf, \quad \dot{x}(t) = u(t), \quad |u(t)| \le 1, \quad x(0) = 0.$$

Тук $x \in AC[0,1]$.

Решение. Щом $x \in AC[0,1]$ и $x(0) = 0$, то

$$x(t) = \int_0^t \dot{x}(\tau)\,d\tau = \int_0^t u(\tau)\,d\tau. \tag{13.1}$$

Следователно, разменяйки реда на интегриране (теорема на Фубини), получаваме

$$\begin{aligned}\int_0^1 x(t)\,dt &= \int_0^1 \left(\int_0^t u(\tau)\,d\tau\right) dt = \int_0^1 u(\tau)\left(\int_\tau^1 dt\right) d\tau \\ &= \int_0^1 (1-\tau)\,u(\tau)\,d\tau.\end{aligned}$$

При условието $|u(t)| \le 1$, последният интеграл горе очевидно достига своя минимум за функцията $u(\tau) = -1$. Сега от (13.1) следва, че решението на задачата е $\widetilde{x}(t) = -t$, при това тази функция е точка на глобален минимум.

Задача 2. Решете екстремалната задача

$$\int_0^1 \left(\dot{x}^2(t) + x^2(t)\right) dt - x^2(1) \longrightarrow \inf, \quad x(0) = 1.$$

Тук $x \in AC[0,1]$.

Решение. Тази задача може да се реши с общата теорема за задачата на Болца (вж. Задача 2 в глава 3) или да се сведе към задача със свободен край, към която да се приложи или твърдението в Задача 1 в глава 3, или Принципът за максимума. Далеч по-лесно е обаче в тази ситуация да разсъждаваме по следния начин.

След като

$$\begin{aligned} x^2(1) &= x^2(0) + \int_0^1 2x(t)\dot{x}(t)\,dt \\ &= 1 + \int_0^1 2x(t)\dot{x}(t)\,dt, \end{aligned}$$

то

$$\int_0^1 \left(\dot{x}^2(t) + x^2(t)\right) dt - x^2(1) = \int_0^1 \left(\dot{x}(t) - x(t)\right)^2 \, dt - 1.$$

Сега вече е очевидно, че екстремалната задача има глобално решение (в силен смисъл) и то е решението на задачата на Коши $\dot{x}(t) = x(t)$, $x(0) = 1$, което е $\widetilde{x}(t) := e^t$.

Задача 3. Изведете Принципа за максимума във форма на Хамилтън за задачата на Майер

$$\left| \begin{aligned} &\psi(t_1, x(t_1)) \longrightarrow \inf \\ &\dot{x}(t) = \varphi(t, x(t), u(t)), \quad t \in [t_0, t_1], \\ &u(t) \in U, \quad t \in [t_0, t_1], \\ &x(t_0) = x_0, \end{aligned} \right. \tag{13.2}$$

като $t_0 \in \mathbb{R}$ и $x_0 \in \mathbb{R}^n$ са фиксирани. Предполагаме, че фазовата траектория $x(t) = (x_1(t), x_2(t), \dots, x_n(t))$ е векторно-значна функция, чиито координатни функции са абсолютно непрекъснати, допустимото управление $u(t) = (u_1(t), u_2(t), \dots, u_m(t))$ е с координатни функции, които са ограничени и измерими (по Лебег), $U \subseteq \mathbb{R}^m$, функцията $\psi(t, x)$ има непрекъснати частни производни до втори ред в някое отворено множество в $\mathbb{R}^{n+1}$ и изображението $\varphi(t, x, u)$ е непрекъснато по (t, x) и непрекъснато диференцируемо по $t, x_1, \dots, x_n$ в някое отворено множество в $\mathbb{R}^{n+m+1}$, както и ограничено върху неговите ограничени подмножества.

Решение. Без ограничение на общността можем да считаме, че $\psi(t_0, x(t_0)) = 0$. Полагаме

$$f(t, x, u) := \psi'_t(t, x) + \psi'_x(t, x) \cdot \varphi(t, x, u).$$

Тогава задачата (13.2) е еквивалентна на задачата

$$\left| \begin{aligned} &\int_{t_0}^{t_1} f(t, x(t), u(t))\,dt \longrightarrow \inf \\ &\dot{x}(t) = \varphi(t, x(t), u(t)), \quad t \in [t_0, t_1], \\ &u(t) \in U, \quad t \in [t_0, t_1], \\ &x(t_0) = x_0, \end{aligned} \right.$$

където t_0 е фиксирано, а t_1 е свободно. Към нея прилагаме Принципа за максимума в Теорема 9.8 с разменени роли на t_0 и t_1 и $b_1(t,x)$, заместено с $b_0(t,x) := (t-t_0, x-x_0)$ (срв. с Теорема 9.7). Функцията на Понтрягин има вида

$$H(t,x,u,p,\lambda) := p \cdot \varphi - \lambda f = (p - \lambda \psi'_x(t,x)) \cdot \varphi(t,x,u) - \lambda \psi'_t(t,x).$$

Задача 4. Изведете необходимо условие за решението на следната задача на Майер

$$\left| \begin{array}{l} \psi(t_1, x(t_1)) \longrightarrow \inf \\ \dot{x}_i(t) \geq 0 \text{ п.н. в } [t_0, t_1],\ i = 1, \dots, n, \\ x(t_0) = x_0, \end{array} \right. \tag{13.3}$$

като $t_0 \in \mathbb{R}$ и $x_0 \in \mathbb{R}^n$ са фиксирани. Предполагаме, че фазовата траектория $x(t) = (x_1(t), x_2(t), \dots, x_n(t))$ е векторно-значна функция, чиито координатни функции са абсолютно непрекъснати, и функцията $\psi(t,x)$ има непрекъснати частни производни до втори ред в някое отворено множество в $\mathbb{R}^{n+1}$.

Решение. Без ограничение на общността можем да считаме, че $\psi(t_0, x(t_0)) = 0$. Полагаме

$$f(t,x,u) := \psi'_t(t,x) + \psi'_x(t,x) \cdot u.$$

Тогава задачата (13.3) е еквивалентна на задачата

$$\left| \begin{array}{l} \displaystyle\int_{t_0}^{t_1} f(t, x(t), u(t))\, dt \longrightarrow \inf \\ \dot{x}(t) = u(t), \quad t \in [t_0, t_1], \\ u(t) \in [0, +\infty)^n, \\ x(t_0) = x_0, \end{array} \right. \tag{13.4}$$

където t_0 е фиксирано, а t_1 е свободно. Към нея прилагаме Принципа за максимума в Теорема 9.8 (срв. с предходната задача) с $\varphi(t,x,u) := u$, $U = [0, +\infty)^n$ и $b_0(t,x) := (t - t_0, x - x_0)$, като $m = n$ и $s_0 = n + 1$.

Функцията на Понтрягин има вида

$$H(t,x,u,p,\lambda) := p \cdot \varphi - \lambda f = (p - \lambda \psi'_x(t,x)) \cdot u - \lambda \psi'_t(t,x).$$

Полагаме и

$$\mathcal{H}(t,x,p,\lambda) := \sup_{u \in [0,+\infty)^n} H(t,x,u,p,\lambda).$$

Ако $(\widetilde{x}(t), \widetilde{u}(t), [t_0, \widetilde{t}_1])$ е решение на (13.4), то според Теорема 9.8(i) съществуват $\lambda \geq 0$, $\ell = (\ell_0, \ell_1, \dots, \ell_n) \in \mathbb{R}^{n+1}$ и $p \in AC([t_0, \widetilde{t}_1], \mathbb{R}^n)$ такива, че

$$\begin{aligned} &\dot{p}(t) = -H'_x(t, \widetilde{x}(t), \widetilde{u}(t), p(t), \lambda) = \lambda \frac{d}{dt} \psi'_x(t, \widetilde{x}(t)) \text{ п.н. в } [t_0, \widetilde{t}_1], \\ &p(t_0) = (\ell_1, \dots, \ell_n), \\ &p(\widetilde{t}_1) = 0. \end{aligned} \tag{13.5}$$

Диференцианото уравнение и второто гранично условие влекат

$$p(t) = \lambda[\psi'_x(t, \widetilde{x}(t)) - \psi'_x(\widetilde{t}_1, \widetilde{x}(\widetilde{t}_1))], \quad t \in [t_0, \widetilde{t}_1]. \tag{13.6}$$

Според Теорема 9.8(ii) $H(t, \widetilde{x}(t), u, p(t), \lambda)$ достига най-голяма стойност по $u \in [0, +\infty)^n$ в $\widetilde{u}(t)$ за почти всяко $t \in [t_0, \widetilde{t}_1]$. Оттук установяваме, че

$$[p_i(t) - \lambda \psi'_{x_i}(t, \widetilde{x}(t))]\widetilde{u}_i(t) = 0 \text{ п.н. в } [t_0, \widetilde{t}_1], \quad i = 1, \dots, n, \tag{13.7}$$

и

$$\mathcal{H}(t, \widetilde{x}(t), p(t), \lambda) = -\lambda \psi'_t(t, \widetilde{x}(t)), \quad t \in [t_0, \widetilde{t}_1]. \tag{13.8}$$

За второто равенство горе сме взели още предвид, че $\mathcal{H}(t, \widetilde{x}(t), p(t), \lambda)$ и $\psi'_t(t, x)$ са непрекъснати (Теорема 9.8(iii)).

От Теорема 9.8(iii) следва, че

$$\mathcal{H}(t_0, \widetilde{x}(t_0), p(t_0), \lambda) = -\ell_0 \tag{13.9}$$

$$\mathcal{H}(\widetilde{t}_1, \widetilde{x}(\widetilde{t}_1), p(\widetilde{t}_1), \lambda) = 0. \tag{13.10}$$

Предвид равенството (13.8), условието (13.10) означава, че

$$\lambda \psi'_t(\widetilde{t}_1, \widetilde{x}(\widetilde{t}_1)) = 0. \tag{13.11}$$

Според Теорема 9.8 множителите λ, ℓ и p не са едновременно нула. Ако допуснем, че $\lambda = 0$, от (13.6) и (13.5) получаваме, че $p(t) \equiv \mathbf{0}$ и $\ell_i = 0$, $i = 1, \dots, n$. Също така от (13.8) следва, че $\mathcal{H}(t, \widetilde{x}(t), p(t), \lambda) \equiv 0$, което влече заедно с (13.9), че и $\ell_0 = 0$. Достигнахме до противоречие. Следователно $\lambda \neq 0$. Тогава от (13.11) следва, че

$$\psi'_t(\widetilde{t}_1, \widetilde{x}(\widetilde{t}_1)) = 0,$$

а от (13.7), (13.6) и $\widetilde{u}(t) = \dot{\widetilde{x}}(t)$ п.н. следва, че

$$\psi'_{x_i}(\widetilde{t}_1, \widetilde{x}(\widetilde{t}_1))\, \dot{\widetilde{x}}_i(t) = 0 \text{ п.н. в } [t_0, \widetilde{t}_1], \quad i = 1, \dots, n.$$

За $i \in \{1, \dots, n\}$ това условие е еквивалентно на

$$\psi'_{x_i}(\widetilde{t}_1, \widetilde{x}(\widetilde{t}_1)) = 0 \quad \text{или} \quad \widetilde{x}_i(t) \equiv x_{0,i},$$

където $x_0 := (x_{0,1}, \dots, x_{0,n})$.

Задача 5. Решете екстремалната задача

$$\left| \begin{array}{l} T \longrightarrow \inf \\ \ddot{x}(t) = u(t), \quad t \in [0,T], \\ |u(t)| \le 1, \quad t \in [0,T], \\ x(0) = \xi, \quad \dot{x}(0) = 0, \\ x(T) = \dot{x}(T) = 0. \end{array} \right.$$

Решение. Задачата може да се представи в стандартния вид за задачи от оптималното управление:

$$\left| \begin{array}{l} \displaystyle\int_0^T dt \longrightarrow \inf \\ \dot{x}_1(t) = x_2(t) \quad t \in [0,T], \\ \dot{x}_2(t) = u(t) \quad t \in [0,T], \\ u(t) \in [-1,1], \quad t \in [0,T], \\ x_1(0) = \xi,\ x_2(0) = 0, \\ x_1(T) = 0,\ x_2(T) = 0. \end{array} \right.$$

Тук фазовите траектории удовлетворяват стандартното предположение $x_1, x_2 \in AC[0,T]$, а управлението $u(t)$ е ограничена измерима функция. От първото диференциално уравнение следва, че дори имаме $x_1 \in C^1[0,T]$.

Прилагаме Теорема 9.3 с $f(t,x_1,x_2,u) := 1$, $\varphi(t,x_1,x_2,u) := (x_2,u)$, $U := [-1,1]$, $b_0(t,x_1,x_2) := (t, x_1 - \xi, x_2)$ и $b_1(t,x_1,x_2) := (x_1,x_2)$, като $n = 2$, $m = 1$, $s_0 = 3$ и $s_1 = 2$. Функцията на Понтрягин има вида

$$H(t,x_1,x_2,u,p_1,p_2,\lambda) := p_1 x_2 + p_2 u - \lambda.$$

Пишем (x_1,x_2) вместо x и (p_1,p_2) вместо p.

Нека $(\widetilde{x}_1(t), \widetilde{x}_2(t), \widetilde{u}(t), [0,\widetilde{T}])$ е решение на задачата. От Теорема 9.3 следва, че съществуват $\lambda \ge 0$, $\ell_0 := (\alpha,\beta,\gamma) \in \mathbb{R}^3$, $\ell_1 := (\delta,\epsilon) \in \mathbb{R}^2$ и $p_1, p_2 \in AC[0,\widetilde{T}]$ такива, че удовлетворяват спрегнатото уравнение в Теорема 9.3(i), което се свежда до

$$\left| \begin{array}{l} \dot{p}_1 = 0, \\ \dot{p}_2 = -p_1 \end{array} \right. \tag{13.12}$$

и граничните условия (условията за трансверзалност)

$$p_1(0) = \beta, \quad p_1(\widetilde{T}) = -\delta,$$
$$p_2(0) = \gamma, \quad p_2(\widetilde{T}) = -\epsilon.$$

Оттук веднага получаваме

$$p_1(t) = \beta = -\delta, \tag{13.13}$$
$$p_2(t) = -\beta\, t + \gamma, \tag{13.14}$$
$$\beta\, \widetilde{T} - \gamma = \epsilon. \tag{13.15}$$

Ако допуснем, че $p_2(t) \equiv 0$, т.е. ако $\beta = \gamma = 0$, то от (13.13) и (13.15) следва, че и $\delta = \epsilon = 0$. Следователно $H(t, x_1, x_2, u, p_1, p_2, \lambda) = -\lambda$, откъдето $\mathcal{H}(t, \widetilde{x}_1(t), \widetilde{x}_2(t), p_1(t), p_2(t), \lambda) = -\lambda$ за всяко t. Сега от Теорема 9.3(iii) следва, че

$$\lambda = \alpha, \quad \lambda = 0.$$

Така всички множители на Лагранж: $\lambda, p_1, p_2, \ell_1$ и ℓ_2 се оказаха равни на нула, което е противоречие. Така установихме, че $p_2(t) \not\equiv 0$.[1]

По-натък от Теорема 9.3(ii) следва, че

$$\widetilde{u}(t) = \operatorname{sgn} p_2(t) = \operatorname{sgn}(\gamma - \beta\, t) \quad \text{п.н. в} \quad [0, \widetilde{T}],$$

оттук и диференциалното уравнение на задачата следва, че

$$\ddot{\widetilde{x}}(t) = \operatorname{sgn}(\gamma - \beta\, t) \quad \text{п.н. в} \quad [0, \widetilde{T}], \tag{13.16}$$

като β и γ не са едновременно нула. Оттук ще определим $\widetilde{x}$. След това ще намерим $\widetilde{T}$, като най-малкото положително число, което едновеременно удовлетворява условията

$$\widetilde{x}(\widetilde{T}) = \dot{\widetilde{x}}(\widetilde{T}) = 0,$$

стига да има такова.

Ако допуснем, че $\beta = 0$, то $\gamma \neq 0$ и с $c := \operatorname{sgn}\gamma$ от (13.16) получаваме $\dot{\widetilde{x}}(t) = c\,t + c_1$. Сега от граничните условия $\dot{\widetilde{x}}(0) = \dot{\widetilde{x}}(\widetilde{T}) = 0$ следва $c = 0$, което е противоречие. Следователно $\beta \neq 0$.

Полагаме $t_0 := \gamma/\beta$. Ако $t_0 \notin (0, \widetilde{T})$, то $\operatorname{sgn}(\gamma - \beta\, t)$ приема една и съща стойност, $+1$ или -1, в целия интервал $(0, \widetilde{T})$ и отново получаваме, че $\dot{\widetilde{x}}(t)$ е линейна функция с ненулев коефициент пред t, което противоречи

[1] В съгласие с направените коментари в част 9.4 нека споменем, че за да достигнем до това заключение, може и да не въвеждаме множителите на Лагранж ℓ_0 и ℓ_1, а само да разгледаме p_1, p_2 и λ, за които знаем, че не са едновременно нула (вж. Теорема 9.7 с $t_0 := 0$ и Теорема 9.5, но модифицирана за фиксирана стойност на $x(t) := (x_1(t), x_2(t))$ в десния край на интервала). Отново започваме със спрегнатото уравнение (13.12). След това отново допускаме, че $p_2(t) \equiv 0$. Това влече $p_1(t) \equiv 0$. Отново достигаме до $H(t, x_1, x_2, u, p_1, p_2, \lambda) = -\lambda$, откъдето $\mathcal{H}(t, \widetilde{x}_1(t), \widetilde{x}_2(t), p_1(t), p_2(t), \lambda) = -\lambda$ за всяко t. Сега от аналога на Теорема 9.5(iii) за фиксирана стойност на $x(t)$ в десния край следва и $\lambda = 0$.

на граничните условия $\dot{\widetilde{x}}(0) = \dot{\widetilde{x}}(\widetilde{T}) = 0$. Следователно $t_0 \in (0, \widetilde{T})$ и от (13.16) имаме п.н. в $[0, \widetilde{T}]$ с $c := \operatorname{sgn}\gamma$

$$\ddot{\widetilde{x}}(t) = \operatorname{sgn}(\gamma - \beta\, t) = \begin{cases} c, & t \in [0, t_0), \\ -c, & t \in (t_0, \widetilde{T}]. \end{cases}$$

Като интегрираме и вземем предвид условието $\dot{\widetilde{x}}(0) = 0$, получаваме

$$\dot{\widetilde{x}}(t) = \begin{cases} ct, & t \in [0, t_0], \\ c(2\, t_0 - t), & t \in [t_0, \widetilde{T}]. \end{cases}$$

От $\dot{\widetilde{x}}(\widetilde{T}) = 0$ следва $2\, t_0 = \widetilde{T}$. Така имаме

$$\dot{\widetilde{x}}(t) = \begin{cases} ct, & t \in [0, t_0], \\ c(\widetilde{T} - t), & t \in [t_0, \widetilde{T}]. \end{cases}$$

Накрая, като интегрираме още веднъж и вземем предвид, че $\widetilde{x}(0) = \xi$, достигаме до

$$\widetilde{x}(t) = \begin{cases} \xi + \dfrac{ct^2}{2}, & t \in [0, \widetilde{T}/2], \\ \xi + \dfrac{c\widetilde{T}^2}{4} - \dfrac{c(\widetilde{T} - t)^2}{2}, & t \in [\widetilde{T}/2, \widetilde{T}]. \end{cases} \tag{13.17}$$

От условието $\widetilde{x}(\widetilde{T}) = 0$ получаваме $\widetilde{T}^2 = -4c\xi$. Ако $\xi = 0$, то решението е $\widetilde{T} = 0$ (този случай е тривиален). Иначе, ако екстремалната задача има решение, то $c = -\operatorname{sgn}\xi$ и $\widetilde{T} = 2\sqrt{|\xi|}$. Така остава да покажем, че $\widetilde{T} = 2\sqrt{|\xi|}$ е действително решение на задачата.

Ще разгледаме подробно случая $\xi < 0$. Случаят $\xi > 0$ е симетричен относно абцисната ос. Оптималната фазова траектория, ако съществува, има вида

$$\widetilde{x}(t) = \begin{cases} \dfrac{t^2}{2} - \dfrac{\widetilde{T}^2}{4}, & t \in [0, \widetilde{T}/2], \\ -\dfrac{(\widetilde{T} - t)^2}{2}, & t \in [\widetilde{T}/2, \widetilde{T}], \end{cases}$$

а оптималното управление е (п.н.)

$$\widetilde{u}(t) = \begin{cases} 1, & t \in [0, \widetilde{T}/2], \\ -1, & t \in (\widetilde{T}/2, \widetilde{T}]. \end{cases}$$

Забележете, че $\widetilde{u}(t)$ удовлетворява ограничението върху управленията като равенство.

Представени посредством даденото те имат вида

$$\widetilde{x}(t) = \begin{cases} \dfrac{t^2}{2} - |\xi|, & t \in \left[0, \sqrt{|\xi|}\right], \\ -\dfrac{(2\sqrt{|\xi|} - t)^2}{2}, & t \in \left[\sqrt{|\xi|}, 2\sqrt{|\xi|}\right], \end{cases}$$

и

$$\widetilde{u}(t) = \begin{cases} 1, & t \in \left[0, \sqrt{|\xi|}\right], \\ -1, & t \in \left(\sqrt{|\xi|}, 2\sqrt{|\xi|}\right]. \end{cases}$$

Ако $x, \dot{x} \in AC[0,T]$, $\ddot{x}(t) = u(t)$, където $u(t)$ е ограничена измерима функция, и $x(0) = \xi$, $\dot{x}(0) = x(T) = \dot{x}(T) = 0$, то за $t \in [0,T]$ имаме

$$\dot{x}(t) = \int_0^t u(\tau)\, d\tau, \tag{13.18}$$

$$\dot{x}(t) = -\int_t^T u(\tau)\, d\tau, \tag{13.19}$$

$$x(t) = \xi + \int_0^t \dot{x}(\tau)\, d\tau. \tag{13.20}$$

Аналогично за $t \in [0, \widetilde{T}]$ имаме

$$\dot{\widetilde{x}}(t) = \int_0^t \widetilde{u}(\tau)\, d\tau, \tag{13.21}$$

$$\dot{\widetilde{x}}(t) = -\int_t^{\widetilde{T}} \widetilde{u}(\tau)\, d\tau, \tag{13.22}$$

$$\widetilde{x}(t) = \xi + \int_0^t \dot{\widetilde{x}}(\tau)\, d\tau. \tag{13.23}$$

Нека $\widetilde{T}/2 < T < \widetilde{T}$. Ако $|u(t)| \le 1$ п.н. в $[0,T]$, то

$$u(t) \le \widetilde{u}(t) \quad \text{п.н. в} \quad [0, \widetilde{T}/2] \tag{13.24}$$

$$\widetilde{u}(t) \le u(t) \quad \text{п.н. в} \quad [\widetilde{T}/2, T]. \tag{13.25}$$

От (13.18), (13.21) и (13.24) следва, че за $t \in [0, \widetilde{T}/2]$ имаме

$$\dot{x}(t) = \int_0^t u(\tau)\, d\tau \le \int_0^t \widetilde{u}(\tau)\, d\tau = \dot{\widetilde{x}}(t);$$

аналогично от (13.19), (13.22) и (13.25) следва, че за $t \in [\widetilde{T}/2, T]$ имаме

$$\dot{x}(t) = -\int_t^T u(\tau)\, d\tau \le -\int_t^T \widetilde{u}(\tau)\, d\tau \le -\int_t^{\widetilde{T}} \widetilde{u}(\tau)\, d\tau = \dot{\widetilde{x}}(t).$$

Така установихме, че

$$\dot{x}(t) \le \dot{\widetilde{x}}(t), \quad x \in [0, T].$$

От това неравенство, (13.20) и (13.23) получаваме аналогично, че

$$x(t) \le \widetilde{x}(t), \quad x \in [0, T];$$

в частност $0 = x(T) \le \widetilde{x}(T)$. Но от друга страна, функцията $\widetilde{x}(t)$ е строго монотонно растяща и, след като $T < \widetilde{T}$, а $\widetilde{x}(\widetilde{T}) = 0$, то $\widetilde{x}(T) < 0$ — противоречие.

Случаят $T \in (0, \widetilde{T}/2]$ се разглежда аналогично. Отново, ако $|u(t)| \le 1$ п.н. в $[0, T]$, то

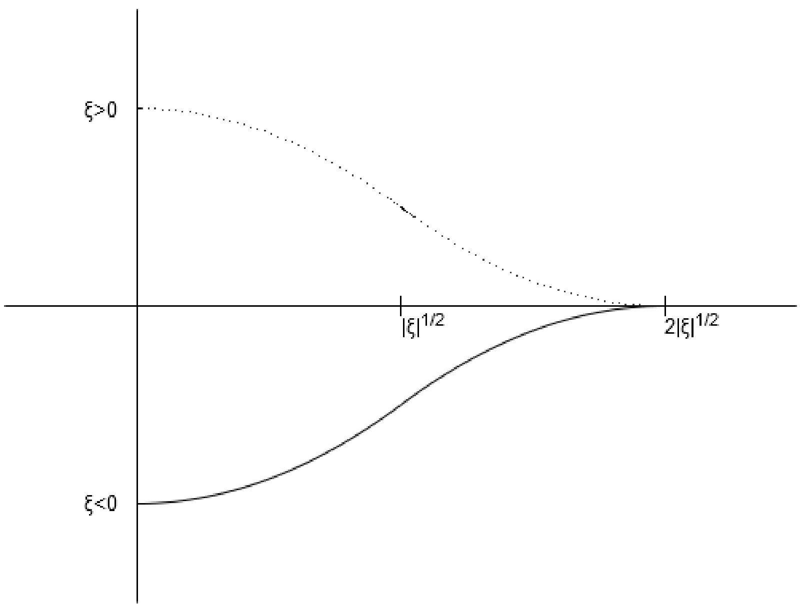
$$u(t) \le \widetilde{u}(t) \quad \text{п.н. в} \quad [0, T],$$

откъдето отново получваме, че

$$x(t) \le \widetilde{x}(t), \quad t \in [0, T],$$

и достигаме до противоречие.

Така установихме, че няма фазова траектория и съответно управление, удовлетворяващи условията на задачата, за които $T < \widetilde{T}$. С това оптималността на процеса $(\widetilde{x}(t), \widetilde{u}(t), [0, \widetilde{T}])$ е доказана. Фигура 3 представя графиките на оптималната фазова траектория в двата случая $\xi < 0$ и $\xi > 0$.



Фигура 3: Оптималната фазова траектория в Задача 5.

Следващите задачи са за самостоятелна работа.

Задача 6. Решете екстремалните задачи, като навсякъде $x = x(t) \in \mathbb{R}$:

а) $\displaystyle\int_0^T \left(\dot{x}^2 + x^2\right) dt \longrightarrow \inf,\ x(0) = 0,\ x(T) = \xi,\ T$ е фиксирано;

б) $\displaystyle\int_0^T \left(\dot{x}^2 - x^2\right) dt \longrightarrow \inf,\ x(0) = 0,\ x(T) = \xi,\ T$ е фиксирано;

в) $\displaystyle\int_{t_0}^{t_1} \sqrt{1 + \dot{x}^2}\, dt \longrightarrow \inf,\ x(t_i) = x_i,\ i = 0, 1,\ [t_0, t_1]$ е фиксиран;

г) $\displaystyle\int_{t_0}^{t_1} x\sqrt{1 + \dot{x}^2}\, dt \longrightarrow \inf,\ x(t_i) = x_i,\ i = 0, 1,\ [t_0, t_1]$ е фиксиран;

д) $\displaystyle\int_0^T \left(\ddot{x}^2 + x^2\right) dt \longrightarrow \inf,\ x(0) = \dot{x}(0) = 0,\ x(T) = \xi_1,\ \dot{x}(T) = \xi_2,\ T$ е фиксирано;

е) $\displaystyle\int_0^1 x^2\, dt \longrightarrow \sup,\ \dot{x} = u,\ |u| \leq 1,\ x(0) = 0;$

ж) $\displaystyle\int_0^1 x\, dt \longrightarrow \sup,\ |\dot{x}| \leq 1,\ x \leq A,\ x(0) = x(1) = 0.$

з) $T \longrightarrow \inf,\ \ddot{x} = u,\ |u| \leq 1,\ \dot{x} \geq -1,\ x(0) = \dot{x}(0) = 0,\ x(T) = \xi_1,\ \dot{x}(t) = \xi_2.$

Задача 7. Намерете точката, сумата на разстоянията от която до четири дадени точки в една равнина е минимална.

Задача 8. Намерете точката, сумата на разстоянията от която до четири дадени точки е минимална.

Задача 9. Намерете най-късата равнинна крива, свързваща две дадени точки, удовлетворяваща условието, че разстоянието между кривата и началото (нулата в равнината) трябва да бъде поне дадено положително число.

Използвана литература

1. Т. Боянов, Записки по теория на екстремалните задачи, 1996.

2. Д. Дойчинов, Математически анализ в крайно-мерни пространства, Университетско издателство "Св. Климент Охридски", София, 1992.

3. А. Йофе, В. Тихомиров, Теория экстремаьлных задач, Издателство „Наука", Москва, 1974; англ. превод: A. D. Ioffe and V. M. Tihomirov, Theory of Extremal Problems, North-Holland Publishing Company, 1979.

4. А. Н. Колмогоров, С. В. Фомин, Элементы теории функций и функционального анализа, Издателство "Наука", Москва, 1976.