ГОСУДАРСТВЕННЫЙ КОМИТЕТ РОССИЙСКОЙ ФЕДЕРАЦИИ
ПО ВЫСШЕМУ ОБРАЗОВАНИЮ

НОВОСИБИРСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
Механико-математический факультет

В.Н.Старовойтов

ПРАВИЛО МНОЖИТЕЛЕЙ ЛАГРАНЖА

Методическое пособие

Новосибирск
1995

Данное пособие педназначается для студентов 2-го, 3-го курсов. При изучении вариационного исчисления необходимо знать некоторые факты и понятия из функционального анализа, чего нельзя требовать от студентов младших курсов. В данном пособии изложены первоначальные сведения о классическом вариационном исчислении, причем, от читателя требуется лишь знание математического анализа в рамках первых двух курсов университета. Достаточно даже знания одной только теоремы о неявной функции.

Часть 1. Необходимое условие экстремума.

§1. Функционалы в линейном нормированном пространстве.

Напомним определение некоторых понятий, которые будут постоянно использоваться в дальнейшем изложении.

Определение. Непустое множество L элементов x, y, z называется *линейным пространством*, если оно удовлетворяет следующим условиям:

1. для любых двух элементов $x, y \in L$ однозначно определяется третий элемент $z \in L$, называемый их суммой и обозначаемый $x + y$, причем

   а) $x + y = y + x$,

   б) $x + (y + z) = (x + y) + z$,

   в) существует такой элемент $0 \in L$, что $x + 0 = x$ для любого $x \in L$,

   г) для любого $x \in L$ существует элемент $(-x) \in L$ такой, что $x + (-x) = 0$.

2. для любых чисел $\alpha, \beta \in R$ и любого элемента $x \in L$ определены элементы $\alpha x,\ \beta x \in L$, причем,

   а) $\alpha(\beta x) = (\alpha\beta)x$,

   б) $1 \cdot x = x$,

   в) $(\alpha + \beta)x = \alpha x + \beta x$,

   г) $\alpha(x + y) = \alpha x + \alpha y$.

Пример.

1. Пространство векторов R^n.

2. $C(a, b)$ — пространство функций, непрерывных на отрезке $[a, b] \subset R$

3. Пусть $M = \{u \in C(a, b) | u(a) = 1\}$. Множество M не является линейным пространством. В самом деле, пусть $u, v \in M$, тогда $(u + v)(a) = u(a) + v(a) = 2$, т. е. $(u + v) \notin M$.

Определение. Линейное пространство E называется *нормированным* (л.н.п.), если каждому $x \in E$ поставлено в соответствие действительное число $\|x\|$, называемое *нормой* элемента x, причем,

1. $\|x\| \geq 0$ для любого $x \in E$,

2. $\|x\| = 0$ тогда и только тогда, когда $x = 0$,

3. $\|\alpha x\| = |\alpha| \|x\|$,

4. $\|x + y\| \leq \|x\| + \|y\|$

для любых $x, y \in E$, $\alpha \in R$.
Пример.

1. В R^n можно ввести норму как длину вектора

$$\|\bar{x}\| = \sqrt{x_1^2 + ... + x_n^2},$$

   и оно превратится в л.н.п.

2. $C(a,b)$ — л.н.п. с нормой $\|u\| = \max\limits_{t\in[a,b]} |u(t)|$.

3. $C^k(a,b)$ — л.н.п. функций, непрерывных на отрезке $[a,b]$ вместе со своими производными до k-го порядка включительно. Норма в этом пространстве
$$\|u\| = \max_{t\in[a,b]} \{|u(t)|, |u'(t)|, ..., |u^{(k)}(t)|\}.$$

Часто, чтобы отличить норму одного пространства от нормы другого, знак нормы снабжают индексом — обозначением соответствующего пространства: $\|\cdot\|_E$.

Определение. Числовую функцию, определенную на некотором л.п., будем называть *функционалом.*

Мы будем рассматривать функционалы над л.н.п.

Определение. Функционал F над л.н.п. E называется *непрерывным в точке* $x_0 \in E$, если для любой последовательности $x_n \in E$, сходящейся к x_0 (т.е. $\|x_n - x_0\| \to 0$ при $n \to \infty$) $F(x_n) \to F(x_0)$.

Если F непрерывен в каждой точке пространства E, то говорят, что F — *непрерывный функционал* на E.

Пример.

1. Норма является непрерывным функционалом.

2. $E = C(a,b)$, $F_1(u) = \int\limits_a^b u(t)\,dt$, $F_2(u) = u(a)$. F_1, F_2 — непрерывные функционалы над $C(a,b)$.

Определение. Функционал F над л.п. E называется *линейным*, если он:

а) аддитивен: $F(x+y) = F(x) + F(y), \quad x, y \in E$,

б) однороден: $F(\alpha x) = \alpha F(x), \quad x \in E, \alpha \in R$.

Выше приведенные функционалы F_1, F_2 являются линейными.

Множество линейных непрерывных функционалов над л.н.п. E образует л.н.п., называемое *сопряженным пространством* и обозначаемое E^*. Норма в нем

$$\|F\|_* = \sup_{\|x\|=1} |F(x)|.$$

Действие линейного функционала L на x часто будем обозначать так: $L\langle x\rangle$.

Отметим некоторые свойства линейных функционалов.

1)$|F\langle x\rangle| \leq \|F\|_* \|x\|$ *для любых* $x \in E, F \in E^*$.

Действительно, так как $|F\langle x\rangle| \leq \|F\|_*$ $\forall x \in E$ такого, что $\|x\| = 1$, то для $\forall y \in E$ справедливо

$$|F\langle y\rangle| = \|y\| |F\langle y/\|y\|\rangle| \leq \|y\| \|F\|_*,$$

что и требовалось доказать.

2)*Для того, чтобы линейный функционал F над л.н.п. E был непрерывен, необходимо и достаточно, чтобы он был ограничен, т.е.* $\|F\|_* \leq K$ *для некоторой константы* K.

Доказательство. Достаточность. Пусть $\|F\|_* \leq K$, x_0 — произвольный элемент из E и x_n — произвольная последовательность, сходящаяся к x_0. Тогда в силу свойства 1

$$|F\langle x_n\rangle - F\langle x_0\rangle| = |F\langle x_n - x_0\rangle| \leq \|F\|_* \|x_n - x_0\| \to 0$$

при $n \to \infty$, т.е. F — непрерывен.

Необходимость. Пусть F — непрерывный линейный функционал. Предположим, что он неограничен, т.е. $\|F\|_* = \infty$.

Это значит, что $\forall n$ существует x_n такой, что $\|x_n\| = 1$, $|F\langle x_n\rangle| > n$. Но тогда при $n \to \infty$ последовательность $y_n = x_n/n$ стремится к нулю, т.к. $\|y_n\| = 1/n \to 0$. В то же время $|F\langle y_n\rangle| > 1$ для любого n, т.е. $|F\langle y_n\rangle|$ к нулю не стремится. Это противоречит непрерывности F, т.к. линейный функционал в нуле равен нулю.

§2. Дифференцирование в линейном пространстве.

Пусть E — л.н.п. Функционал F над E называется *дифференцируемым* в точке $x \in E$, если существует такой линейный ограниченный функционал L_x, что

$$F(x+h) - F(x) - L_x\langle h\rangle = o(h)$$

для любого $h \in E$. Величина $o(h)$ такова, что

$$\lim_{\|h\|\to 0} \frac{o(h)}{\|h\|} = 0$$

Линейный функционал L_x называется *сильной производной* или *производной Фреше* функционала F в точке x и обозначается $F'(x)$.

Покажем, что производная определена единственным образом. Предположим, что существуют два непрерывных линейных функционала L_1, L_2, которые являются производными функционала F в точке x, т.е.

$$F(x+h) - F(x) - L_i\langle h\rangle = o(h), \quad i = 1, 2,$$

тогда $L_1\langle h\rangle - L_2\langle h\rangle = o(h)$ для любого $h \in E$. Это значит, что

$$\lim_{\|h\|\to 0} \frac{|L_1\langle h\rangle - L_2\langle h\rangle|}{\|h\|} = 0.$$

Возьмем произвольный элемент $\varphi \in E$ такой, что $\|\varphi\| = 1$. Тогда

$$0 = \lim_{\lambda\to 0} \frac{|L_1\langle \lambda\varphi\rangle - L_2\langle \lambda\varphi\rangle|}{\|\lambda\varphi\|} = \lim_{\lambda\to 0} \frac{|\lambda||L_1\langle \varphi\rangle - L_2\langle \varphi\rangle|}{|\lambda|\|\varphi\|} =$$

$$= |L_1\langle\varphi\rangle - L_2\langle\varphi\rangle|.$$

Отсюда следует, что $\|L_1 - L_2\|_* = 0$, т.е. $L_1 = L_2$. Можно более подробно: $L_1\langle\varphi\rangle = L_2\langle\varphi\rangle$ для любого $\varphi \in E$ такого, что $\|\varphi\| = 1$. Следовательно, для любого $h \in E$

$$L_1\langle h\rangle = \|h\| L_1\langle h/\|h\|\rangle = \|h\| L_2\langle h/\|h\|\rangle = L_2\langle h\rangle.$$

Т.е. $L_1 = L_2$.

УПРАЖНЕНИЕ. Доказать, что дифференцируемый функционал непрерывен.

Посчитаем производные некоторых функционалов.

1. Пусть $F(x) = \alpha = const$. Тогда $F'(x) = 0$ для любого $x \in E$. В самом деле, нулевой функционал является производной функционала F:

   $$F(x+h) - F(x) - 0 = 0 = o(h),$$

   а производная определена единственным образом.

2. Пусть F — линейный ограниченный функционал, тогда

   $$F'(x)\langle h\rangle = F\langle h\rangle.$$

   Это следует из того, что $F(x+h) - F(x) - F\langle h\rangle = 0 = o(h)$ в силу линейности F.

3. Пусть F, G — дифференцируемые функционалы над E, тогда $F+G, \alpha F$ — дифференцируемы и

$$(F+G)'(x) = F'(x) + G'(x), \;\; (\alpha F)'(x) = \alpha F'(x).$$

Действительно,

$$(F+G)'(x)\langle h\rangle = (F+G)(x+h) - (F+G)(x) + o(h) =$$
$$= F(x+h) - F(x) + G(x+h) - G(x) + o(h) =$$
$$= F'(x)\langle h\rangle + G'(x)\langle h\rangle + o(h),$$

откуда следует первое равенство. Второе доказывается аналогично.

Пример. Пусть $E = C(a,b), F(u) = \int\limits_a^b u^2(t)dt$. Тогда

$$F'(u)\langle h\rangle = 2\int\limits_a^b u(t)h(t)dt.$$

В самом деле,

$$F(u+h) - F(u) = \int\limits_a^b [(u(t)+h(t))^2 - u^2(t)]dt =$$

$$= \int\limits_a^b 2u(t)h(t) + h^2(t)dt.$$

Покажем, что

$$\int\limits_a^b h^2(t)dt = o(h).$$

Действительно,

$$\int\limits_a^b h^2(t)dt \le \max_{t\in[a,b]} |h(t)|^2(b-a) = \|h\|^2(b-a).$$

Следовательно,

$$\lim_{\|h\|\to 0} \frac{\int\limits_a^b h^2(t)dt}{\|h\|} \le \lim_{\|h\|\to 0}(b-a)\|h\| = 0,$$

что и требовалось доказать.

Определение. Пусть F — функционал над л.н.п. E и $x_0 \in E$. Если функция $F(x_0+\lambda h)$ переменного λ дифференцируема при $\lambda = 0$ для любого $h \in E$, то ее производная

$$\frac{d}{d\lambda}F(x_0+\lambda h)\Big|_{\lambda=0} = \lim_{\lambda\to 0}\frac{F(x_0+\lambda h)-F(x_0)}{\lambda}$$

называется *вариацией* функционала F в точке x_0 и обозначается $\delta F(x_0,h)$.

Вариация может не быть линейной по h. Если же такая линейность имеет место, т.е.

$$\delta F(x_0,h) = F'_c(x_0)\langle h\rangle,$$

где $F'_c(x_0)$ — непрерывный линейный функционал, то этот функционал называется *слабой производной* или *производной Гато* функционала F в точке x_0, а функционал F — *дифференцируемым по Гато* в этой точке.

Пример. Посчитаем слабую производную функционала из предыдущего примера.

$$F_c'(u)\langle h\rangle = \frac{d}{d\lambda}\int\limits_a^b (u(t)+\lambda h(t))^2 dt\Big|_{\lambda=0} =$$

$$= \int\limits_a^b 2(u(t)+\lambda h(t))h(t)\Big|_{\lambda=0} dt = 2\int\limits_a^b u(t)h(t)dt.$$

Видим, что $F_c'(u)\langle h\rangle = F'(u)\langle h\rangle$, хотя в общем случае сильная и слабая дифференцируемость представляют собой разные понятия.

Предложение 2.1. *Если функционал F имеет сильную производную, то он имеет и слабую, причем сильная и слабая производные совпадают.*

Доказательство. Пусть F сильно дифференцируем в точке x. Тогда

$$F(x+th) - F(x) = F'(x)\langle th\rangle + o(th) = tF'(x)\langle h\rangle + o(th).$$

Следовательно,

$$\delta F(x,h) = \lim_{t\to 0}\frac{F(x+th)-F(x)}{t} =$$

$$= F'(x)\langle h\rangle + \lim_{t\to 0}\frac{o(th)}{t} = F'(x)\langle h\rangle.$$

т.е. вариация является линейным по h функционалом, поэтому $F_c'(x)\langle h\rangle = \delta F(x,h)$ а следовательно,

$$F_c'(x)\langle h\rangle = F'(x)\langle h\rangle.$$

Обратное утверждение, вообще говоря, неверно, из слабой дифференцируемости не следует сильная. Можно даже привести пример разрывной функции, у которой тем не менее существует слабая производная. Сильная производная в этом случае, конечно же, существовать не может. *Пример.* Пусть $E = R^2, x = (x_1, x_2)$,

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

Очевидно, что в точке $(0,0)$ функция F разрывна. Посчитаем вариацию в этой точке. Видим, что $\delta F(0,h) = 0$ для любого h, т.е. производная Гато существует и равна нулю:

$$F_c'(0) = 0.$$

Это связано с тем, что когда мы считаем вариацию, мы стремимся к точке 0 вдоль прямых, проходящих через эту точку. А вдоль прямых эта функция непрерывна в точке 0 и равна нулю.

Таким образом, вариация $\delta F(x,h)$ есть производная по направлению h в точке x.

Выясним условия, при которых из слабой дифференцируемости функционала F следует его сильная дифференцируемость.

Определение. Множество $S_r(x_0) = \{x \in E | \|x - x_0\| < r\}$ называется *открытым шаром* радиуса r с центром в точке x_0. $B_r(x_0) = \{x \in E | \|x - x_0\| \le r\}$ — *замкнутый шар.*

Предложение 2.2. *Если слабая производная $F_c'(x)$ функционала F существует в некотором шаре $S_r(x_0)$ и непрерывна по x в точке x_0, т.е.*

$$\lim_{\|y\| \to 0} \|F_c'(x_0 + y) - F_c'(x_0)\|_* = 0,$$

то в точке x_0 *сильная производная существует и совпадает со слабой.*

Доказательство. Пусть $h \in E$ — произвольная точка, такая, что $\|h\| < r$, т.е. $(x_0 + h) \in S_r(x_0)$. Рассмотрим функцию действительного переменного $\varphi(t) = F(x_0 + th)$. По условию эта функция дифференцируема. Тогда по теореме Лагранжа существует $\theta \in [0, 1]$ такое, что

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

То есть

$$F(x_0 + h) - F(x_0) = \frac{d}{dt}F(x_0 + th)\Big|_{t=\theta} =$$

$$= \frac{d}{d\lambda}F((x_0 + \theta h) + \lambda h)\Big|_{\lambda=0} = F_c'(x_0 + \theta h)\langle h\rangle.$$

Здесь $\lambda = t - \theta, (x_0 + \theta h) \in S_r(x_0)$.

Теперь,

$$F(x_0 + h) - F(x_0) - F_c'(x_0)\langle h\rangle =$$

$$= F_c'(x_0 + \theta h)\langle h\rangle - F_c'(x_0)\langle h\rangle = o(h).$$

В самом деле,

$$\lim_{\|h\|\to 0} \frac{|F_c'(x_0 + \theta h)\langle h\rangle - F_c'(x_0)\langle h\rangle|}{\|h\|} \le$$

$$\le \lim_{\|h\|\to 0} \frac{\|F_c'(x_0 + \theta h) - F_c'(x_0)\|_* \|h\|}{\|h\|} = 0$$

Таким образом,

$$F_c'(x_0)\langle h\rangle = F'(x_0)\langle h\rangle.$$

что и требовалось доказать.

§3. Необходимое условие экстремума.

Определение. Говорят, что функционал F над E имеет в точке x_0 *относительный* или *локальный минимум (максимум)*, если существует такой шар $S_r(x_0)$, что для всех $x \in S_r(x_0)$

$$F(x_0) \leq F(x) \quad (F(x_0) \geq F(x)).$$

А если $F(x_0) \leq F(x)$ $(F(x_0) \geq F(x))$ для любого $x \in E$, то x_0 — точка *абсолютного минимума (максимума)* функционала F. Очевидно, что точка абсолютного минимума (максимума) является также точкой относительного минимума (максимума).

Определение. Точка x_0 является точкой *экстремума* функционала F, если эта точка — либо точка относительного минимума, либо относительного максимума.

Основная задача вариационного исчисления — находить точки, в которых заданный функционал достигает своего экстремального значения.

Теорема 3.1. *Пусть $x_0 \in E$ является точкой экстремума функционала F, и в этой точке существует вариация $\delta F(x_0, h)$, тогда $\delta F(x_0, h) = 0$ для любого $h \in E$.*

Доказательство. Рассмотрим функцию действительного переменного $\varphi(t) = F(x_0 + th)$. В точке $t = 0$ эта функция достигает экстремума, поэтому $\varphi'(t)|_{t=0} = 0$. То есть

$$\delta F(x_0, h) = \frac{d}{dt} F(x_0 + th)\Big|_{t=0} = 0.$$

Здесь h было произвольным. Теорема доказана.

Определение. Точка x_0 называется *стационарной точкой* функционала F, если в этой точке вариация $\delta F(x_0, h)$ определена и $\delta F(x_0, h) = 0$ для любого $h \in E$.

Теорема 3.1 дает только необходимое условие экстремума. Очевидно, что если функционал F дифференцируем, то необходимым условием экстремума является $F'(x_0) = 0$, так как в этом случае вариация совпадает с сильной производной.

§4. Правило множителей Лагранжа.

В предыдущем параграфе мы рассматривали задачу о нахождении экстремума функционала во всем пространстве, но чаще всего приходится иметь дело с задачами, в которых есть те или иные ограничения. Например, задача о нахождении кривой заданной длины, ограничивающей наибольшую площадь. Заданность длины кривой — это и есть ограничение. Сформулируем задачу математически.

Пусть $\Phi_i,\ i = 1, ..., N$ — дифференцируемые функционалы на л.н.п. E и $M = \{x \in E | \Phi_i(x) = 0, i = 1, ..., N\}$.

Определение. Будем говорить, что функционал F достигает в точке $x_0 \in M$ экстремума при условиях $\Phi_i(x) = 0, i = 1, ..., N$, если существует шар $S_r(x_0) \subset E$ такой, что

$$F(x_0) \leq F(x) \text{ для любого } x \in M \cap S_r(x_0),$$

либо

$$F(x_0) \geq F(x) \text{ для любого } x \in M \cap S_r(x_0).$$

Сначала докажем ряд вспомогательных утверждений. Предположим, что у нас имеется конечный набор линейных функционалов $L_i, i = 1, ..., N$, причем, они линейно независимы. Кроме того, ни одно из уравнений $L_i\langle x\rangle = 0$ не удовлетворяется тождественно для всех $x \in E$. Тогда множество $E' = \{x \in E | L_i\langle x\rangle = 0, i = 1, ..., N\}$ является линейным подпространством в E.

Теорема 4.1. *Всякий элемент x из E может быть представлен в виде*

$$x = \sum_{i=1}^{N} t_i a_i + a,$$

где $a_1, ...a_n$ есть система фиксированных линейно независимых элементов, лежащих вне E', элемент a принадлежит E'.

Доказательство. Проведем доказательство индукцией по N. Пусть $N = 1$ и подпространство E' определяется единственным уравнением $L\langle x\rangle = 0$. В силу наших оговорок существует элемент $a_1 \in E$ такой, что $L\langle a_1\rangle \neq 0$. Пусть x — произвольный элемент из E. Обозначим

$$a = x - \frac{L\langle x\rangle}{L\langle a_1\rangle} a_1$$

Очевидно, $L\langle a\rangle = 0$, т.е. $a \in E'$ и $x = a + ta_1$, где $t = L\langle x\rangle / L\langle a_1\rangle$. Первый шаг индукции доказан.

Предположим, что теорема верна для случая $N - 1$ уравнений, определяющих подпространство E', и докажем ее для случая N уравнений.

Пусть E' — линейное подпространство, определенное N уравнениями: $L_i\langle x\rangle = 0, i = 1, ...N$, а E_N есть линейное подпространство, определенное одним уравнением $L_N\langle x\rangle = 0$. В силу только что доказанного для произвольного x из E имеем

$$x = b + ta_N$$

где $L_N\langle b\rangle = 0, L_N\langle a_N\rangle \neq 0$.

Теперь заметим, что E_N само является линейным пространством, и E' — его линейное подпространство, определенное в E_N $(N-1)$ уравнениями: $L_i\langle x\rangle = 0, i = 1, ..., N-1$, так как N-е уравнение в E_N выполняется тождественно.

Так как по предположению теорема доказана для случая систем из $(N-1)$ уравнений, то на E_N существуют $(n-1)$ линейно независимых элементов $a_1, ..., a_{N-1}$, лежащих вне E' таких, что для произвольного элемента из E_N, а следовательно, для b из E_N имеем

$$b = a + \sum_{i=1}^{N-1} t_i a_i,$$

где $a \subset E'$. Таким образом,

$$x = a + \sum_{i=1}^{N} t_i a_i.$$

Обратив внимание, что a_N лежит вне E_N, а $a_i, i = 1, ..., N-1$ лежат на E_N и что $a_1, ..., a_{N-1}$ между собой линейно независимы, получим, что все $a_1, ..., a_N$ линейно независимы. Теорема доказана.

Теорема 4.2. *Если линейное подпространство* E'

$$L_i\langle \bar{x}\rangle = 0, \quad i = 1, ..., N,$$

принадлежит линейному подпространству E_0*:*

$$L_0\langle x\rangle = 0,$$

то существует такой набор постоянных $\lambda_1, ..., \lambda_N$*, что*

$$L_0\langle x\rangle + \sum_{i=1}^{N} \lambda_i L_i\langle x\rangle \equiv 0.$$

Доказательство. По предыдущей теореме, всякий элемент представим в виде

$$x = a' + \sum_{i=1}^{N} t_i a_i,$$

где $a' \in E'$ а значит, $a' \in E_0, a_1, ..., a_N$ линейно независимы и лежат вне E'. Поэтому

$$L_j\langle x\rangle = \sum_{i=1}^{N} t_i L_j\langle a_i\rangle, \quad j = 1, ..., N,$$

$$L_0\langle x\rangle = \sum_{i=1}^{N} t_i L_0\langle a_i\rangle.$$

В силу линейной независимости элементов $a_i, i = 1, ..., N$ и функционалов $L_i, i = 1, ..., N$, определитель матрицы $A_{ij} = (L_j(a_i)), i, j = 1, ..., N$, отличен от нуля. Поэтому, разрешая первую систему относительно t_i и подставляя решение во вторую систему, получим

$$L_0\langle x\rangle + \sum_{k=1}^{N} \lambda_k L_k\langle x\rangle = 0,$$

где λ_k выражаются через $L_i\langle a_j\rangle$, $L_0\langle a_j\rangle$.

Теорема доказана.

Сейчас мы докажем теорему Л.Люстерника, которая является основным моментом в выводе правила множителей Лагранжа.

Итак, пусть $\Phi_i, i = 1, ..., n$ — дифференцируемые функционалы над E, $M = \{x \in E | \Phi_i(x) = 0, i = 1, ..., n\}$.

Определение. Точка $x_0 \in M$ называется неособой, если все линейные функционалы $\Phi_i'(x_0)$ линейно независимы и

$$\Phi_i'(x_0) \not\equiv 0 \quad i = 1, ..., n.$$

Замечание. Если $n \geq 2$, то второе условие в этом определении следует из первого. В самом деле, если какое-то $\Phi_k'(x_0) \equiv 0$, т.е. $\Phi_k'(x_0)\langle h\rangle = 0 \quad \forall h \in E$, то можно взять набор чисел $\lambda_1 = 0, ..., \lambda_k = 1, ..., \lambda_n = 0$ такой, что $\sum \lambda_i \Phi_i'(x_0) \equiv 0$. А это противоречит первому условию.

Если же $n = 1$, то в этом определении "работает"второе условие, а первое теряет смысл.

Теорема 4.3 (Люстерник). *Пусть $x_0 \in M$ — неособая точка, $h \in E' = \{g \in E | \Phi_i'(x_0)\langle g\rangle = 0, i = 1, ..., n\}$. Тогда существуют n линейно независимых элементов $h_1, ..., h_n$, лежащих вне E', а также n функций $s_1, ..., s_n$ действительного аргумента таких, что*
1. точка $x = x_0 + th + \sum\limits_{i=1}^{n} h_i s_i(t)$ принадлежит M при достаточно малых по абсолютной величине значениях t,
2. $s_i(0) = 0, \frac{ds_i}{dt}\Big|_{t=0} = 0, i = 1, ..., n.$

Доказательство. Согласно теореме 4.1, для любого $x \in E$ справедливо представление

$$x - x_0 = a + \sum_{i=1}^{n} s_i h_i,$$

где $a \in E'$, $h_i \notin E', i = 1, ..., n$. Возьмем $h \in E'$ и рассмотрим только те $x \in E$, которые имеют вид

$$x = x_0 + th + \sum_{i=1}^{n} s_i h_i,$$

при этом будем менять t в окрестности нуля. Элемент $x \in M$, если $\Phi_j(x) = 0, j = 1, ..., n$, то есть

$$u_j(t, s_1, ..., s_n) \equiv \Phi_j(x_0 + th + \sum_{i=1}^{n} s_i h_i) = 0, \quad j = 1, ..., n.$$

Для доказательства существования требуемых функций s_i воспользуемся теоремой о неявной функции. Пусть матрица A имеет элементы

$$A_{ij} = \frac{\partial u_i}{\partial s_j}.$$

Тогда $A_{ij}\Big|_{t=0, s_k=0} = \Phi_i'(x_0)\langle h_j \rangle$, поэтому $\det A\Big|_{t=0, s_k=0} \neq 0$ в силу линейной независимости h_j и функционалов $\Phi_i'(x_0)$. По теореме о неявной функции систему уравнений

$$u_j(t, s_1, ..., s_n) = 0, \quad j = 1, ..., n$$

можно разрешить в явном виде в некоторой окрестности точки $t = 0, s_i = 0, i = 1, ..., n$, которая, естественно, удовлетворяет системе, т.е. $x_0 \in M$. Таким образом, существуют функции $s_1(t), ..., s_n(t)$ такие, что для любого t из некоторой окрестности нуля наша система уравнений выполняется. При этом $s_i(0) = 0, i = 1, ..., n$,

$$\frac{ds_i}{dt}\Big|_{t=0} = -\sum_{j=1}^{n} \frac{\partial u_i}{\partial t} B_{ij}\Big|_{t=0, s_i=0}, i = 1, ..., n,$$

где $B = A^{-1}$. Так как

$$\frac{\partial u}{\partial t}\Big|_{t=0, s_i=0} = \Phi_j'(x_0)\langle h \rangle = 0, j = 1, ..., n,$$

в силу того, что $h \in E'$, то

$$\frac{ds_i}{dt}\Big|_{t=0} = 0.$$

Теорема доказана.

Теперь мы можем перейти к выводу необходимых условий экстремума функционала при наличии ограничений.

Теорема 4.4 (правило множителей Лагранжа). *Пусть функционалы $F, \Phi_1, ..., \Phi_n$ дифференцируемы в некоторой окрестности точки x_0, удовлетворяющей условиям $\Phi_i(x_0) = 0$, $i = 1, ..., n$. Если в точке x_0 достигается локальный экстремум функции $F(x)$ на множестве $\{x|\Phi_i(x) = 0, i = 1, ..., n\}$, то существуют числа $\lambda_0, \lambda_1, ..., \lambda_n$, не все равные нулю, такие, что*

$$\lambda_0 F'(x_0) + \sum_{i=1}^{n} \lambda_i \Phi_i'(x_0) \equiv 0,$$

т.е.

$$\lambda_0 F'(x_0)\langle h\rangle + \sum_{i=1}^{n} \lambda_i \Phi_i'(x_0)\langle h\rangle = 0 \quad \forall h \in E.$$

Доказательство. Если x_0 — особая точка, то $\Phi_i'(x_0), i = 1, ..., n$ — линейно зависимы. Т.е. существует набор чисел $\lambda_1, ..., \lambda_n$, среди которых есть ненулевые, такой, что

$$\sum_{i=1}^{n} \lambda_i \Phi_i'(x_0) \equiv 0.$$

Взяв теперь $\lambda_0 = 0$, получим утверждение теоремы.

Пусть теперь x_0 — неособая точка. Применим теорему Люстерника. Возьмем $h \in E'$ и $s_i(t), i = 1, ..., n$ такие, что точка $x = x_0 + th + \sum\limits_{i=1}^{n} h_i s_i(t)$ принадлежит M при достаточно малых по модулю t.

Рассмотрим функцию

$$v(t) = F(x_0 + th + \sum_{i=1}^{n} h_i s_i(t)).$$

Очевидно, что эта функция имеет экстремум при $t = 0$, поэтому

$$\frac{dv}{dt}\,|_{t=0} = 0.|$$

Но

$$\frac{dv}{dt}\Big|_{t=0} = F'(x_0)\langle h\rangle + F'(x_0)\langle h_i\rangle \frac{ds_i}{dt}\Big|_{t=0} = F'(x_0)\langle h\rangle.$$

То есть $F'(x_0)\langle h\rangle = 0$. Таким образом, любой элемент подпространства $E' = \{y \in E | \Phi_i'(x_0) = 0, i = 1, ..., n\}$ принадлежит подпространству, задаваемому уравнением $F'(x_0)\langle y\rangle = 0$. По теореме 4.2 существует такой набор постоянных $\lambda_1, ..., \lambda_n$, что

$$F'(x_0) + \sum_{i=1}^{n} \lambda_i \Phi_i'(x_0) \equiv 0,$$

то есть утверждение теоремы справедливо при $\lambda_0 = 1$. Теорема доказана.

Рассмотрим теперь несколько примеров.

Пример.

1. Задача без ограничений.

Пусть

$$E = C(0,1), F(u) = \int\limits_0^1 (u(t) - t^2)^2 \, dt,$$

$$F'(u)\langle h\rangle = 2\int\limits_0^1 (u(t) - t^2)h(t) \, dt.$$

Лемма (Лагранж). Если функция $f(t)$ $(a \leq t \leq b)$ непрерывна и для любой непрерывной функции $h(t)$ выполняется условие

$$\int\limits_a^b f(t)h(t)dt = 0,$$

то $f(t) = 0$.

Эта лемма будет доказана в следующем парагафе. Применяя эту лемму к нашему случаю, получим, что функция $u(t) \in C(0,1)$ доставляет минимум функционалу F только в том случае, если $u(t) = t^2$. Легко убедиться непосредственной проверкой, что действительно функция t^2 является точкой минимума функционала F.

2. Задача с ограничением. Пусть $E = C(0,1)$,

$$F(u) = \int\limits_0^1 (u(t) - t^2)dt.$$

Будем искать минимум функционала F на множестве непрерывных функций, удовлетворяющих условию

$$\int\limits_0^1 u(t)dt = 1.$$

Применим правило множителей Лагранжа. В нашем случае $n = 1$, $\Phi(u) \equiv \int\limits_0^1 u(t)dt - 1$. Очевидно, что $\Phi'(u) \not\equiv 0$ при любом u, поэтому особых точек нет. Тогда

$$F'(u)\langle h\rangle + \lambda\Phi'(u)\langle h\rangle = \int\limits_0^1 [2(u(t) - t^2) + \lambda]h(t)dt = 0.$$

По лемме Лагранжа

$$2(u(t) - t^2) + \lambda = 0,$$

т.е. $u(t) = t^2 - \lambda/2$. Но λ нам неизвестно. Оно находится из условия $\Phi(u) = 0$:

$$\int\limits_0^1 (t^2 - \frac{\lambda}{2})dt = 1.$$

Отсюда $\lambda = -4/3$. Таким образом, искомая функция

$$u(t) = t^2 + 2/3.$$

Это стационарная точка функционала F при заданном условии. Является ли она точкой локального минимума или максимума, надо проверять дополнительно.

Часть 2. Одномерные вариационные задачи.

§5. Уравнения Эйлера-Лагранжа.

В этой части мы будем заниматься вариационными задачами с функционалами, определенными на пространстве функций, зависящих только от одной переменной. Такие задачи мы назовем одномерными. Чаще всего нам будут встречаться задачи на экстремум для функционалов вида

$$\int\limits_a^b L(t, \bar{x}(t), \dot{\bar{x}}(t))dt,$$

где $\bar{x} : R \to R^n$ — вектор-функция действительного переменного t, $\dot{\bar{x}}$ — производная функции $\bar{x}$ по t. Подынтегральную функцию L называют *лагранжианом* или *интегрантом.*

Прежде чем переходить к выводу необходимых условий экстремума указанных функционалов, докажем утверждение, которое в дальнейшем будет постоянно использоваться.

Лемма Дюбуа — Реймона. *Пусть на отрезке* $[a, b]$ *функции* $f_0(t)$ *и* $f_1(t)$ *непрерывны, и пусть для любой непрерывной дифференцируемой функции* $h(t)$, *для которой* $h(a) = h(b) = 0$, *выполнено равенство*

$$\int\limits_a^b (f_1(t)\dot{h}(t) + f_0(t)h(t))dt = 0.$$

Тогда функция f_1 непрерывно дифференцируема и

$$-\frac{df_1}{dt} + f_0 = 0.$$

Доказательство. Возьмем функцию $p \in C^1(a,b)$ такую, что $\dot{p}(t) = f_0(t)$ и $\int\limits_a^b p(t)dt = \int\limits_a^b f_1(t)dt$. Тогда для любой функции $h \in C^1(a,b)$, для которой $h(a) = h(b) = 0$, по условию леммы должно выполняться равенство

$$0 = \int\limits_a^b (f_1(t)\dot{h}(t) + f_0(t)h(t))dt = \int\limits_a^b (f_1(t)\dot{h}(t) + \dot{p}(t)h(t))dt =$$

$$= \int\limits_a^b (f_1(t) - p(t))\dot{h}(t)dt.$$

Выберем функцию $g \in C^1(a,b)$ такую, что

$$\dot{g}(t) = f_1(t) - p(t), \quad g(a) = 0.$$

Тогда

$$g(b) = \int\limits_a^b \dot{g}(t)dt = \int\limits_a^b (f_1(t) - p(t))dt = 0$$

в силу выбора функции p.

Положим $h(t) = g(t)$. Тогда

$$\int\limits_a^b (f_1(t) - p(t))^2 dt = 0,$$

т.е. $p(t) = f_1(t)$. Значит, $f_1 \in C^1(a, b)$,

$$-\frac{df_1}{dt} + f_0 = 0.$$

Следствие (лемма Лагранжа). *Пусть $f \in C(a, b)$ и для любой непрерывной функции h такой, что $h(a) = h(b) = 0$, выполняется равенство*

$$\int_a^b f(t)h(t)dt = 0,$$

то $f(t) \equiv 0$.

Рассмотрим следующую задачу: среди всех функций $y(x) : [a, b] \to R$, принадлежащих $C^1(a, b)$ и удовлетворяющих условиям

$$y(a) = A, \quad y(b) = B,$$

найти такую, которая доставляет минимум (экстремум) функционалу

$$F(y) = \int_a^b L(x, y, \frac{dy}{dx})dx,$$

где $L(x, y, dy/dx)$ — дважды непрерывно дифференцируемая функция своих аргументов. Мы выведем только необходимые условия того, чтобы функция $y(x)$ являлась решениями этой задачи. Так как это задача с ограничениями, воспользуемся правилом множителей Лагранжа.

Пусть $\Phi_1(y) = y(a) - A$, $\Phi_2(y) = y(b) - B$. Это линейные функционалы над $C^1(a, b)$. Очевидно, они дифференцируемы. Для того, чтобы функция $y(x)$ доставляла локальный условный экстремум функционалу F, необходимо, чтобы существовал набор чисел $\lambda_0, \lambda_1, \lambda_2$ таких, что

$$\lambda_0 F'(y) + \lambda_1 \Phi_1'(y) + \lambda_2 \Phi_2'(y) = 0.$$

Так как функционалы $\Phi_1'(y), \Phi_2'(y)$ линейно независимы, можно взять $\lambda_0 = 1$. Далее,

$$F'(y)\langle h \rangle = \int_a^b \left(\frac{\partial L}{\partial y} h + \frac{\partial L}{\partial y_x} \frac{dh}{dx} \right) dx,$$

$$\Phi_1'(y)\langle h \rangle = h(a), \quad \Phi_2'(y)\langle h \rangle = h(b).$$

Получаем

$$\int_a^b \left(\frac{\partial L}{\partial y} h + \frac{\partial L}{\partial y_x} \frac{dh}{dx} \right) dx + \lambda_1 h(a) + \lambda_2 h(b) = 0, \qquad (*)$$

для любой функции $h \in C^1(a,b)$, а значит, и для всех тех $h \in C^1(a,b)$, которые обращаются в нуль в точках $x = a$ и $x = b$. Тогда по лемме Дюбуа — Реймона функция $\partial L / \partial y_x$ непрерывно дифференцируема как функция от x, и

$$\frac{d}{dx} \left(\frac{\partial L}{\partial y_x} \right) - \frac{\partial L}{\partial y} = 0.$$

Это уравнение называется уравнением Эйлера — Лагранжа. Кроме этого уравнения, функция $y(x)$ должна удовлетворять ограничениям $\Phi_i(y) = 0$, то есть

$$y(a) = A, \quad y(b) = B.$$

Можно теперь посчитать и множители Лагранжа, хотя это и не нужно. С помощью интегрирования по частям из (*) получим

$$\int_a^b \left(\frac{\partial L}{\partial y}h - \frac{d}{dx}\frac{\partial L}{\partial y_x}\right) h(x)dx + h(a)\left(\lambda_1 - \frac{\partial L}{\partial y_x}\Big|_{x=a}\right) +$$

$$+h(b)\left(\lambda_2 + \frac{\partial L}{\partial y_x}\Big|_{x=b}\right) = 0.$$

Первое слагаемое равно нулю в силу уравнения Эйлера — Лагранжа. Что касается двух других слагаемых, то так как функция h произвольна, возьмем ее сначала такой, что $h(a) = 1$, $h(b) = 0$, а потом $h(a) = 0$, $h(b) = 1$. Получим

$$\lambda_1 = \frac{\partial L}{\partial y_x}\Big|_{x=a}, \quad \lambda_2 = -\frac{\partial L}{\partial y_x}\Big|_{x=b}.$$

Совершенно аналогично можно получить уравнения Эйлера — Лагранжа в случае векторных функций. Пусть мы имеем вектор-функцию $\bar{y}(x)$: $[a, b] \to R^n$, принадлежащую $C^1(a, b)$ и доставляющую минимум функционалу

$$F(\bar{y}) = \int_a^b L(x, \bar{y}, \bar{y}_x)dx$$

при условиях $\bar{y}(a) = \bar{A}, \bar{y}(b) = \bar{B}$. Тогда $\bar{y}(x)$ должна быть решением задачи

$$\frac{d}{dx}\left(\frac{\partial L}{\partial \bar{y}_x}\right) - \frac{\partial L}{\partial \bar{y}} = 0, \quad \bar{y}(a) = \bar{A}, \quad \bar{y}(b) = \bar{B}.$$

Здесь $\partial L/\partial \bar{y}$ — вектор с компонентами $\partial L/\partial y_i$, $i = 1, ..., n$. Решение уравнения Эйлера — Лагранжа называются экстремалями функционала F.

Исследуем теперь гладкость решения полученной задачи. Дело в том, что мы искали экстремум в пространстве $C^1(a, b)$, а уравнение Эйлера — Лагранжа имеет второй порядок. Конечно, из леммы Дюбуа — Реймона следует, что $\partial L/\partial y_x$ непрерывно дифференцируема по x, но хотелось бы выяснить, определены ли вторые производные у функции $y(x)$.

Теорема Гильберта. *Если $\bar{y}(x)$ есть экстремаль функционала F и если в некоторой точке $x = c \in (a, b)$ имеет место неравенство*

$$det\left(\frac{\partial^2 L}{\partial y_{ix} \partial y_{jx}}\right)_{i,j=1}^{n} \neq 0,$$

то в некоторой окрестности точки $x = c$ вектор-функция $\bar{y}(x)$ имеет непрерывную вторую производную.

Доказательство. Проинтегрируем уравнение Эйлера — Лагранжа от a до произвольной точки $z \in (a, b)$:

$$\frac{\partial L}{\partial \bar{y}_x}(z, \bar{y}(z), \bar{y}'(z)) = \int\limits_a^z \frac{\partial L}{\partial \bar{y}}(x, \bar{y}(x), \bar{y}'(x)) + \bar{p},$$

где $\bar{p} = \ \partial L/\partial \bar{y}_x|_{x=a} = \text{const}$, $\bar{y}'(z) = \bar{y}_z = d\bar{y}/dz$. Напомним, что функция $\bar{y}(x)$ нам уже известна и мы лишь исследуем ее гладкость. Кроме того, $\bar{y}(x) \in C^1(a, b)$, так как мы искали минимум в этом пространстве. Поэтому система уравнений

$$\frac{\partial L}{\partial \bar{y}_x}(z, \bar{y}(z), \bar{q}) = \int\limits_a^z \frac{\partial L}{\partial \bar{y}}(x, \bar{y}(x), \bar{y}'(x))dx + \bar{p}$$

имеет решение для каждого $z \in (a, b)$ ($\bar{q} = \bar{y}'(x)$ является этим решением.) Так как определитель в условии теоремы отличен от нуля в точке $x = c$, то по теореме о неявной функции $\bar{q}(x)$ определяется однозначно этой системой и условием $\bar{q}(c) = \bar{y}'(c)$ в некоторой окрестности точки c и в этой окрестности непрерывно дифференцируема. Таким образом, $\bar{y}''(x)$ существует и непрерывна в окрестности точки $x = c$. Теорема доказана.

Рассмотрим ряд примеров.

Пример 1.

$$E = C^1(0, 1), \quad F(y) = \int\limits_0^1 (\frac{1}{2}x(y')^4 - x^2(y')^2)dx,$$

$$y(0) = 0, \quad y(1) = 2/3.$$

Уравнение Эйлера — Лагранжа:

$$(x(y')^3 - x^2y')' = 0, \Rightarrow x(y')^3 - x^2y' = \text{const}.$$

Так как y' — ограничена, то const=0. Отсюда

$$xy'((y')^2 - x) = 0 \Rightarrow$$

$$\begin{array}{llll} \text{либо} & xy' = 0, & \text{либо} & (y')^2 = x, \\ & y' = 0, & & y' = \pm\sqrt{x}, \\ & y = \text{const}, & & y = \pm\frac{2}{3}x^{3/2} + C, \end{array}$$

Учитывая краевое условие, получим решение

$$y = \frac{2}{3}x^{3/2}.$$

Но $y'' = \frac{1}{2}\frac{1}{\sqrt{x}}$, т.е. при $x = 0$ вторая производная не существует. Это не противоречит теореме Гильберта. В самом деле, $L_{y'y'} = 0$ при $x = 0$. Заметим также, что $L_{y'} = x(y')^3 - x^2y' \equiv 0 \in C^1(0,1)$.

Пример 2. Решение задачи с ограничениями существует не всегда.

$$E = C^1(0,1), \quad F(x) = \int\limits_0^1 t^2\dot{u}^2 dt, \quad u(0) = 0, \quad u(1) = 1.$$

Уравнение Эйлера — Лагранжа:

$$\frac{d}{dt}(2t^2\dot{u}) = 0 \Rightarrow t^2\dot{u} = c \Rightarrow$$

$$u = \frac{c_1}{t} + c_2.$$

Но эта функция не принадлежит $C^1(0,1)$, поэтому в этом постранстве решения задачи нет.

§6. Вариационные задачи в параметрической форме.

В предыдущем параграфе мы рассматривали задачи, в которых было необходимо соединить две точки, например, на плоскости, кривой, уравнение которой задавалось в виде $y = f(x)$. При этом f предполагалась однозначной функцией, и мы вынуждены были ограничиваться линиями, пересекающими прямые, параллельные оси Oy, только в одной точке. В геометрических задачах удобнее пользоваться другим представлением кривых — параметрическим: $x = \varphi(t), y = \psi(t)$. Правда, одна и та же кривая может иметь много параметрических представлений. Если мы хотим рассматривать функционалы от кривых, то необходимо добиться того, чтобы значение этих функционалов не зависело от способа параметризации кривых.

Рассмотрим кривую

$$x = x(t), \quad y = y(t),$$

и на ней дугу, определяемую значениями параметра t, удовлетворяющими неравенству $t_0 \leq t \leq t_1$. Рассмотрим функционал

$$F(\bar{x}) = \int\limits_{t_0}^{t_1} L(t, x, y, \dot{x}, \dot{y})dt.$$

Здесь $\bar{x} = (x, y)$, $\dot{x} = dx/dt$.

Определение. Интегрант (лагранжиан) L называется *параметрическим*, если он не зависит явно от t и

$$L(x, y, k\dot{x}, k\dot{y}) = kL(x, y, \dot{x}, \dot{y})$$

для любого $k \in R$.

Предложение . *Для того, чтобы значение функционала F не зависело от параметризации кривой необходимо и достаточно, чтобы интегрант этого функционала был параметрическим.*

Доказательство. Необходимость. Пусть заданы две параметризации одной и той же кривой: $x = x(\tau), y = y(\tau), \tau_0 \leq \tau \leq \tau_1$; $\quad \tilde{x} = \tilde{x}(t), \tilde{y} = \tilde{y}(t), t_0 \leq t \leq t_1$, причем параметры t и τ связаны соотношениями $t = \chi(\tau),\ \chi'(\tau) > 0$ или < 0.

То, что функционал не зависит от параметризации, запишется так:

$$\int\limits_{\tau_0}^{\tau_1} L(\tau, x, y, \frac{dx}{d\tau}, \frac{dy}{d\tau})d\tau = \int\limits_{t_0}^{t_1} L(t, \tilde{x}, \tilde{y}, \frac{d\tilde{x}}{dt}, \frac{d\tilde{y}}{dt})dt.$$

Продолжим это равенство, сделав замену переменной $t = \chi(\tau)$:

$$= \int\limits_{\tau_0}^{\tau_1} L\left(\chi(\tau), x, y, \frac{dx}{d\tau}\Big/\frac{dt}{d\tau}, \frac{dy}{d\tau}\Big/\frac{dt}{d\tau}\right)\frac{dt}{d\tau}d\tau =$$

$$= \int\limits_{\tau_0}^{\tau_1} L(\chi(\tau), x, y, \frac{dx}{d\tau}\Big/\chi'(\tau), \frac{dy}{d\tau}\Big/\chi'(\tau))\chi'(\tau)d\tau.$$

Если мы потребуем, чтобы независимость интеграла от выбора параметра имела место для любой дуги, то последнее равенство должно выполняться при любом значении верхнего предела τ_1. Продифференцировав его по τ_1 получим:

$$L\left(\chi(\tau), x, y, \frac{x'}{\chi'}, \frac{y'}{\chi'}\right)\chi' = L(\tau, x, y, x', y').$$

Так как χ — произвольная функция, возьмем $\chi(\tau) = \tau + \alpha$, где α - произвольное число. Тогда

$$L(\tau + \alpha, x, y, x', y') = L(\tau, x, y, x', y'),$$

откуда следует, что интегант L не должен явно зависеть от τ.

Тепеь выберем χ так, чтобы $\chi' \equiv 1/k$, k =const. Для любого k

$$kL(x, y, x', y') = L(x, y, kx', ky').$$

Необходимость доказана. Для доказательства достаточности нужно провести все рассуждения в обратном порядке.

Мы провели рассуждения в случае кривых в R^2. В R^n все делается аналогично.

Заметим, что если у нас имеется функционал вида

$$F(y) = \int\limits_a^b f(x, y(x), y'(x))dx,$$

то можно переписать его в параметрической форме. Записав уравнение кривой $y = y(x)$ в виде $x = x(t)$, $y = y(t)$, мы получим функционал с параметрическим интегрантом. В самом деле,

$$\int\limits_a^b f(x, y, y'(x))dx = \int\limits_{t_0}^{t_1} f\left(x(t), y(t), \frac{\dot{y}(t)}{\dot{x}(t)}\right)\dot{x}(t)dt =$$

$$= \int\limits_{t_0}^{t_1} L(x, y, \dot{x}, \dot{y})dt.$$

Очевидно, что L — параметрический интегрант:

$$L(x, y, k\dot{x}, k\dot{y}) = f\left(x, y, \frac{k\dot{y}}{k\dot{x}}\right)k\dot{x} =$$

$$= kf\left(x, y, \frac{\dot{y}}{\dot{x}}\right)\dot{x} = kL(x, y, \dot{x}, \dot{y}).$$

Вообще говоря, любой интегрант можно сделать параметрическим введением новой неизвестной. Пусть

$$F(\bar{x}) = \int\limits_{t_0}^{t_1} L(t, x_1, ..., x_n, \dot{x}_1, ..., \dot{x}_n)dt,$$

$$\bar{x} = (x_1, ..., x_n).$$

Введем новый параметр $\tau : t = x_0(\tau)$, причем функция x_0 неизвестна, лишь $x_0(\tau_0) = t_0$, $x_0(\tau_1) = t_1$. Тогда

$$F(x_0, \bar{x}) = \int\limits_{\tau_0}^{\tau_1} L\left(x_0, ..., x_n, \frac{x_1'}{x_0'}, ..., \frac{x_n'}{x_0'}\right) x_0' d\tau =$$

$$= \int\limits_{\tau_0}^{\tau_1} L_1(x_0, ..., x_n, x_0', ..., x_n')d\tau.$$

L_1 — параметрический интегрант. Здесь $x' = dx/d\tau$.

Пример (задача Дидоны). Исторически одной из первых возникла следующая задача: найти форму замкнутой веревочки заданной длины, такую, чтобы ограничиваемая ею площадь была наибольшей.

Пусть $x = x(t), y = y(t)$ — кривая, задающая форму веревочки. Так как параметризация произвольна, то можно считать, что $t \in [0, 1]$. Возьмем $E = C^1(0, 1)$,

$$F(\bar{x}) = \int\limits_0^1 (x\dot{y} - y\dot{x})dt, \quad \Phi_1(\bar{x}) = \int\limits_0^1 \sqrt{\dot{x}^2 + \dot{y}^2}dt.$$

F задает площадь, Φ_1 — длину дуги. Ограничения: $\Phi_1(\bar{x}) = C$, кроме того, $x(0) = x(1), y(0) = y(1)$, то есть $\Phi_2(\bar{x}) \equiv x(0) - x(1) = 0, \quad \Phi_3(\bar{x}) \equiv y(0) - y(1) = 0$.

Используя правило множителей Лагранжа, получим

$$2\dot{y} - \lambda_1 \frac{d}{dt}\left(\frac{\dot{x}}{\sqrt{\dot{x}^2 + \dot{y}^2}}\right) = 0, \quad 2\dot{x} + \lambda_1 \frac{d}{dx}\left(\frac{\dot{y}}{\sqrt{\dot{x}^2 + \dot{y}^2}}\right) = 0.$$

Эти уравнения нелинейны и решить их довольно трудно. Мы упростим задачу, используя то, что фигурирующие в задаче функционалы имеют пааметрические интегранты.

Так как параметризация произвольна, можно считать, что $\dot{x}^2 + \dot{y}^2 = \mu^{-2}$=const, где μ — некоторая константа. Тогда

$$2\dot{y} - \lambda_1 \mu \ddot{x} = 0, \quad 2\dot{x} + \lambda_1 \mu \ddot{y} = 0.$$

Нам неизвестны λ_1 и μ. Пусть $\lambda = \frac{2}{\lambda_1 \mu}$,

$$\begin{cases} \ddot{x} - \lambda \dot{y} = 0, \\ \ddot{y} + \lambda \dot{x} = 0. \end{cases}$$

Уравнения стали линейными. Решая эту систему, получим:

$$x = \alpha \cos(\lambda t + \varphi) + \beta,$$
$$y = \alpha \sin(\lambda t + \varphi) + \gamma,$$

где $\alpha, \beta, \gamma, \varphi$=const. Так как x, y — периодические, то $\lambda = 2\pi k$, где k — произвольное целое число. Чтобы найти α, надо подставить уравнение кривой в условие $\Phi_1(\bar{x}) = C$. Видим, что найденная нами кривая является окружностью.

Константы φ, β, γ могут быть заданы произвольно. β, γ — координаты центра окужности, φ — фаза, т.е. начало отсчета дуги на окружности. Число k задает количество витков.

§7. Задача с подвижными концами

В §6 была рассмотрена задача о нахождении экстремума функционала в классе кривых, концы которых зафиксированы в заданных точках пространства R^n. Сейчас мы рассмотрим задачу, в которой концы кривых принадлежат некоторым многообразиям. Примером может служить задача о нахождении кратчайшей линии, соединяющей две сферы. Такого рода задачи будем называть задачами с *подвижными концами*. Для простоты мы будем проводить все выкладки в случае плоских кривых. В R^n все делается аналогично.

Итак, пусть заданы две функции $\varphi, \psi \in C^1(R^2)$. Предположим также, что $|\nabla\varphi| \neq 0$, там, где $\varphi = 0$, и $|\nabla\psi| \neq 0$, где $\psi = 0$. Тогда множества $M_1 = \{(x, y) \in R^2 | \varphi(x, y) = 0\}$, $M_2 = \{(x, y) \in R^2 | \psi(x, y) = 0\}$ являются непрерывно дифференцируемыми линиями на плоскости. Задача состоит в нахождении условий, при которых кривая $\{x(t), y(t), t_1 \leq t \leq t_2\}$ доставляет экстремум функционалу

$$F(x, y) = \int_{t_1}^{t_2} L(x(t), y(t), \dot{x}(t), \dot{y}(t))dt$$

с параметрическим лагранжианом L в классе кривых, принадлежащих $C^1(t_1, t_2)$ таких, что

$$(x(t_1), t(t_1)) \in M_1, \quad (x(t_2), y(t_2)) \in M_2.$$

Последние условия можно записать так:

$$\Phi(x, y) \equiv \varphi(x(t_1), y(t_1)) = 0, \quad \Psi(x, y) \equiv \psi(x(t_2), y(t_2)) = 0.$$

Здесь Φ, Ψ являются функционалами.

Заметим, что так как F — функционал с параметрическим лагранжианом, то мы можем менять параметризацию кривых. В частности, можно взять $t_1 = 0$, $t_2 = 1$. Таким образом, мы имеем задачу с ограничениями. Воспользуемся правилом множителей Лагранжа:

$$\lambda_0 F'(x,y)\langle h_1, h_2\rangle + \lambda_1 \Phi'(x,y)\langle h_1, h_2\rangle + \lambda_2 \Psi(x,y)\langle h_1, h_2\rangle = 0,$$

то есть

$$\lambda_0 \int\limits_0^1 \left(\frac{\partial L}{\partial x} h_1 + \frac{\partial L}{\partial y} h_2 + \frac{\partial L}{\partial \dot{x}} \dot{h}_1 + \frac{\partial L}{\partial \dot{y}} \dot{h}_2\right) dt +$$

$$+\lambda_1 \frac{\partial \varphi}{\partial x}\bigg|_{t=0} h_1(0) + \lambda_1 \frac{\partial \varphi}{\partial y}\bigg|_{t=0} h_1(0) +$$

$$+\lambda_2 \frac{\partial \psi}{\partial x}\bigg|_{t=1} h_1(1) + \lambda_2 \frac{\partial \psi}{\partial y}\bigg|_{t=1} h_1(1)$$

для произвольных функций $h_1, h_2 \in C^1(0,1)$. Отсюда, используя лемму Дюбуа — Реймона, с помощью стандартных рассуждений получаем, что решение задачи должно удовлетворять следующим условиям:

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}}\right) - \frac{\partial L}{\partial x} = 0, \quad \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{y}}\right) - \frac{\partial L}{\partial y} = 0 \ \text{ при } \ 0 < t < 1,$$

$$\frac{\partial L}{\partial \dot{x}} = \lambda_1 \frac{\partial \varphi}{\partial x}, \quad \frac{\partial L}{\partial \dot{y}} = \lambda_1 \frac{\partial \varphi}{\partial y} \ \text{ при } \ t = 0,$$

$$\frac{\partial L}{\partial \dot{x}} = -\lambda_2 \frac{\partial \psi}{\partial x}, \quad \frac{\partial L}{\partial \dot{y}} = -\lambda_2 \frac{\partial \psi}{\partial y} \ \text{ при } \ t = 1.$$

Условия при $t = 0$ и $t = 1$ называются *условиями трансверсальности в параметрической форме.* Они выражают тот факт, что вектор $\partial L/\partial \dot{\bar{u}}$, где $\bar{u} = (x, y)$, коллинеарен $\nabla\varphi$ при $t = 0$ и $\nabla\psi$ при $t = 1$. Так как $|\nabla\varphi| \neq 0, |\nabla\psi| \neq 0$, тогда если, например, $\partial\varphi/\partial y \neq 0, \partial\psi/\partial y \neq 0$, то эти условия можно переписать так:

$$\frac{\partial L}{\partial \dot{x}} - \frac{\partial L}{\partial \dot{y}} \frac{\partial \varphi}{\partial x} \Bigg/ \frac{\partial \varphi}{\partial y} = 0 \quad \text{при} \ \ t = 0,$$

$$\frac{\partial L}{\partial \dot{x}} - \frac{\partial L}{\partial \dot{y}} \frac{\partial \psi}{\partial x} \Bigg/ \frac{\partial \psi}{\partial y} = 0 \quad \text{при} \ \ t = 1.$$

Пример. Найти кратчайшую линию, соединяющую две окружности.

$$E = C^1(0, 1), \quad F(x, y) = \int\limits_0^1 \sqrt{\dot{x}^2 + \dot{y}^2} dt,$$

$$\varphi(x, y) = x^2 + y^2 - 1, \quad \psi(x, y) = (x - 4)^2 + y^2 - 1.$$

Используя выведенные формулы, получим, что искомая кривая должна удовлетворять уравнениям

$$\frac{d}{dt}\left(\frac{\dot{x}}{\sqrt{\dot{x}^2 + \dot{y}^2}}\right) = 0, \quad \frac{d}{dt}\left(\frac{\dot{y}}{\sqrt{\dot{x}^2 + \dot{y}^2}}\right) = 0 \ \ \text{при} \ \ 0 < t < 1,$$

а также условиям трансверсальности

$$\frac{\dot{x}}{\sqrt{\dot{x}^2 + \dot{y}^2}} = 2x\lambda_1, \quad \frac{\dot{y}}{\sqrt{\dot{x}^2 + \dot{y}^2}} = 2y\lambda_1 \ \ \text{при} \ \ t = 0,$$

$$\frac{\dot x}{\sqrt{\dot x^2+\dot y^2}} = -2x\lambda_2, \quad \frac{\dot y}{\sqrt{\dot x^2+\dot y^2}} = -2y\lambda_2 \ \text{ при } \ t=1.$$

Так как F — функционал с параметрически интегрантом, то выберем параметр t таким образом, чтобы $\sqrt{\dot x^2+\dot y^2} = \mu =$const, где постоянная μ неизвестна. Тогда задача перепишется так:

$$\ddot x = 0, \quad \ddot y = 0 \ \text{ при } \ t\in(0,1),$$

$$\dot x(0) = \mu_1 x(0), \quad \dot y(0) = \mu_1 y(0),$$
$$\dot x(1) = \mu_2 x(1), \quad \dot y(1) = \mu_2 y(1),$$

где $\mu_1 = 2\mu\lambda_1, \mu_2 = -2\mu\lambda_2$. Сюда надо добавить ограничения

$$\varphi(x(0),y(0)) = x^2(0)+y^2(0)-1=0,$$
$$\psi(x(1),y(1)) = (x(1)-4)^2 - y^2(1) - 1 = 0.$$

Несложно получить, что решением этой задачи будут прямые

$$\text{(а)} \begin{cases} x(t) = 1+2t, \\ y(t)\equiv 0, \end{cases} \quad \text{(б)} \begin{cases} x(t) = -1+6t, \\ y(t)\equiv 0, \end{cases}$$

$$\text{(в)} \begin{cases} x(t) = 1+4t, \\ y(t)\equiv 0, \end{cases} \quad \text{(г)} \begin{cases} x(t) = -1+6t, \\ y(t)\equiv 0. \end{cases}$$

Здесь (а) дает минимум функционала F, (б) — максимум, (в), (г) — стационарные точки.

В данном примере условия трансверсальности означают, что экстремаль подходит под прямым углом к линиям $\varphi = 0$, $\psi = 0$.

Теперь рассмотрим задачу с подвижными концами в непараметрической форме. Предположим, что непрерывно дифференцируемая функция $y(x)$ и два числа x_1^*, x_2^* являются решением следующей вариационной задачи:

$$H(y) = \int_{x_1}^{x_2} G(x, y, y_x) dx = \operatorname{extr} H,$$

$$y(x_1) = f(x_1), \quad y(x_2) = g(x_2),$$

причем x_1, x_2 неизвестны и тоже должны быть найдены. Здесь функции f, g будем считать непрерывно дифференцируемыми.

Используя рассуждения предыдущего параграфа, перейдем к задаче в параметрической форме с параметрическим лагранжианом:

$$F(x, y) = \int_0^1 L(x, y, \dot{x}, \dot{y}) dt,$$

$$\varphi(x, y) \equiv y - f(x) = 0 \quad \text{при} \quad t = 0,$$
$$\psi(x, y) \equiv y - g(x) = 0 \quad \text{при} \quad t = 1,$$

где $L(x, y, \dot{x}, \dot{y}) = \dot{x} G(x, y, \dot{y}/\dot{x})$.

Воспользовавшись формулами, выведенными в начале параграфа, получим уравнения Эйлера — Лагранжа:

$$\frac{d}{dt}\left(G - \frac{\dot{y}}{\dot{x}} \frac{\partial G}{\partial y_x}\right) - \dot{x} \frac{\partial G}{\partial x} = 0,$$

$$\frac{d}{dt}\left(\frac{\partial G}{\partial y_x}\right) - \dot{x} \frac{\partial G}{\partial y} = 0$$

и условия трансверсальности

$$G - \frac{\dot{y}}{\dot{x}} \frac{\partial G}{\partial y_x} = -\lambda_1 \frac{\partial f}{\partial x}, \quad \frac{\partial G}{\partial y_x} = \lambda_1 \text{ при } t = 0,$$

$$G - \frac{\dot{y}}{\dot{x}} \frac{\partial G}{\partial y_x} = \lambda_2 \frac{\partial g}{\partial x}, \quad \frac{\partial G}{\partial y_x} = -\lambda_2 \text{ при } t = 1.$$

В силу произвольности выбора параметра t положим $x = \mu t + \gamma$, где μ, γ — неизвестные постоянные. Мы имеем на это право, так как по предположению решение этой задачи, заданное в явном виде, есть функция $y(x)$. Поэтому $y = y(\mu t + \gamma)$, $x = \mu t + \gamma$ будет параметрическим заданием той же кривой. Таким образом, $\dot{x} = \mu, \dot{y} = y_x \dot{x} = \mu y_x$. Теперь уравнения Эйлера — Лагранжа будут выглядеть так:

$$y_x \left[\frac{d}{dx} \left(\frac{\partial G}{\partial y_x} \right) - \frac{\partial G}{\partial y} \right] = 0,$$

$$\frac{d}{dx} \left(\frac{\partial G}{\partial y_x} \right) - \frac{\partial G}{\partial y} = 0,$$

то есть первое будет следствием второго.

Условия трансверсальности преобразуется к такому виду:

$$G + \frac{\partial G}{\partial y_x} \left(\frac{\partial f}{\partial x} - y_x \right) = 0 \text{ на левом конце,}$$

$$G + \frac{\partial G}{\partial y_x} \left(\frac{\partial g}{\partial x} - y_x \right) = 0 \text{ на правом конце.}$$

Так как на левом конце $x_1 = x_1^*$, а на правом $x_2 = x_2^*$, то $\gamma = x_1^*, \mu + \gamma = x_2^*$.

Итак, нами полностью выписаны условия, которым должно удовлетворять решение задачи с подвижными концами в непараметрической форме.

Литеатура

1.Колмогоров А.Н., Фомин С.В. Элементы теории функций и функционального анализа. М., 1989.
2.Лаврентьев М.А., Люстерник Л.А. Основы вариационного исчисления. 1935. Т.1, ч.2.
3.Гельфанд И.М., Фомин С.В. Вариационное исчисление. М., 1961.

В.Н.Старовойтов

Правило множителей Лагранжа

Методическое пособие

Подписано в печать 02.06.95	Формат $60 \times 84, 1/16$
Офсетная печать, Заказ N	Уч.-изд. л.2,5
Тираж 100 экз.	Цена 800 р.

Редакционно-издательский отдел Новосибирского университета ; участок оперативной полиграфии НГУ; 630090, Новосибирск, ул.Пирогова, 2.