Раздел 8. Экстремальные задачи в ЛНП

Лекция 16 Экстремальные задачи в ЛНП.

Определения экстремумов.

$\Phi : X \to E^1$ – функционал, $M \subset X$ – подмножество.

x_0 – точка локального максимума Φ на M, если $x_0 \in M$ и

$$\exists r > 0\, \forall x \in S_r(x_0) \cap M : \Phi(x) \leq \Phi(x_0)$$

(для минимума $\Phi(x) \geq \Phi(x_0)$).

Если $M = X$ – безусловный экстремум (минимум, максимум), если $M \subsetneq X$ – условный.

Глобальный максимум:

$$\forall x \in M : \Phi(x) \leq \Phi(x_0)$$

(для минимума $\Phi(x) \geq \Phi(x_0)$). Опять может быть условный и безусловный.

В приведённых выражениях неравенства нестрогие, экстремум тоже нестрогий. Если заменить строгими неравенствами при $x \neq x_0$ – будут строгие экстремумы.

Таким образом, экстремумы бывают: максимум и минимум; локальный и глобальный; строгий и нестрогий; условный и безусловный. Любые сочетания.

Как всегда, "точка" = элемент пространства (например, функция).

Замечание. В матанализе различается точка экстремума (эта точка должна быть внутренней) и наибольшее (наименьшее) значение функции. Здесь мы не будем этого различия, и допускается, что x_0 – граничная точка M.

Замечание. Задачи на поиск экстремума (максимума, минимума) называются задачами оптимизации. В частности, различают задачи безусловной и условной оптимизации (поиск безусловного и условного экстремума).

Замечание. Изолированные точки M – точки локального минимума и максимума одновременно. Дальше не рассматриваем.

Замечание. Если M – единичная сфера или единичный замкнутый шар, а функционал Φ линейный, то значение этого функционала в точке максимума (если она существует) – это его норма. Если $\Phi(x) = \|Ax\|$, где A – линейный ограниченный оператор, то то значение такого функционала в точке максимума на M (если она существует) – это норма оператора A.

Необходимое условие экстремума.

Пусть $x(t)$ – абстрактная функция, $x(0) = x_0$, $\forall t \in [0, 1] : x(t) \in M$. Образ функции – кривая в M, начинающаяся в точке x_0. Рассмотрим числовую функцию $\varphi(t) = \Phi(x(t))$. Если x_0 – точка локального максимума (минимума) Φ, то функция φ при $t = 0$ также должна иметь локальный максимум (минимум). Если функция φ имеет одностороннюю производную $\varphi'_+(0)$ при $t = 0$, то эта производная должна быть неположительна (неотрицательна).

Пусть теперь M – выпуклое множество. Тогда в качестве $x(t)$ можно взять функцию x_0+th, где $h=z-x_0$, $z\in M$, образом которой при $t\in[0,1]$ является отрезок, соединяющий x_0 и z и, в силу выпуклости M, целиком лежащий в этом множестве. В этом случае

$$\varphi'_+(0)=\left.\frac{\partial\Phi(x)}{\partial h}\right|_{x=x_0}=\|h\|\,\mathrm{Var}[\Phi(x_0,h^0)]\,,$$

где $h^0=h/\|h\|=(z-x_0)/\|z-x_0\|$ – единичный вектор в направлении $z-x_0$. Отсюда вытекает, что необходимым устовием того, что x_0 является точкой локального максимума (минимума) функционала Φ, является неположительность (неотрицательность) его вариаций вдоль всех допустимых направлений, для которых такие вариации существуют.

Замечание. Допустиме направления – направления, ведущие из x_0 в точки, принадлежащие M.

Если функционал дифференцируем по Фреше в точке x_0, то

$$\left.\frac{\partial\Phi(x)}{\partial h}\right|_{x=x_0}=\Phi'(x_0)(h)=\|h\|\Phi'(x_0)(h^0)\,.$$

Назовём точку x_0 внутренней точкой M, если наряду с произвольным допустимым направлением h^0 также допустимым будет и противоположное ему направление $-h^0$.
(**!!!** Замечание. Это **не соответствует** определению внутренней точки множества, данному ранее. Например, для продпространства, согласно новому определению, все точки будут внутренними, а по старому – все точки граничные, если подпространство не совпадает с X: любая окрестность содержит точки из дополнения.)
В этом случае, поскольку $\Phi'(x_0)(-h)=-\Phi'(x_0)(h)$, мы получаем, что во внутренней точке экстремума

$$\Phi'(x_0)h=0$$

для произвольного допустимого вектора h. Такое условие называется условием стационарности, а точка x_0, в которой оно выполняется – стационарной точкой функционала Φ на M или, если h любое, на всём X.

Замечание. Если $M=X$ (безусловный экстремум), то отсюда следует $\Phi'(x_0)=O$.

Замечание. Если M представляет собою некоторую гиперповерхность в X (невыпуклое, вообще говоря, множество), то h – касательные направления. Множество касательных векторов образует касательное подпространство в пространстве X.

Частный случай: гильбертово пространство. Действие функционала $\Phi'(x_0)$ на элемент h представляется в виде скаларного произведения

$$\Phi'(x_0)(h)=(\mathrm{grad}\,\Phi(x_0),h)\,.$$

Необходимое условие максимума (минимума) – неположительность (неотрицательность) $(\operatorname{grad}\Phi(x_0), h)$ в допустимых направлениях. Для внутренних точек необходимое условие экстремума – равенство $(\operatorname{grad}\Phi(x_0), h) = 0$. В случае безусловного экстремума $\operatorname{grad}\Phi(x_0) = 0$.

Элементы вариационного исчисления.

Пусть $X = C^1[a,b]$, а Φ – интегральный функционал вида

$$\Phi(x) = \int_a^b f(s, x(s), \dot{x}(s))\, ds\,.$$

Классическое вариацационное исчисление занимается, в частности, задачами на поиск максимума или минимума функционалов такого вида.

Замечание. Раньше переменную интегрирования всегда незывали t, но сейчас t у нас занято, поэтому используем s.

Функцию $f(s,u,v)$ будем считать достаточно гладкой. Нам потребуется непрерывность по первой переменной, непрерывная дифференцируемость по второй и третьей, причём частные производные будем считать равномерно липшицевыми:

$$|f'_u(s,u_1,v_1) - f'_u(s,u_2,v_2)| \le L(|u_1-u_2| + |v_1-v_2|)$$
$$|f'_v(s,u_1,v_1) - f'_v(s,u_2,v_2)| \le L(|u_1-u_2| + |v_1-v_2|)$$

(постоянная Липшица L не зависит от s, u и v). В этом случае

$$f(s,u+\xi,v+\eta) - f(s,u,v) = f'_u(s,u,v)\xi + f'_v(s,u,v)\eta + \omega(s,u,v,\xi,\eta)\,,$$

где $|\omega(s,u,v,\xi,\eta)| \le L(|\xi|+|\eta|)^2$ (докажите!)

Замечание. Штрихи при производных будем опускать: $f'_u = f_u$, $f'_v = f_v$.

Замечание. Часто вместо f_u и f_v пишут f_x и $f_{\dot{x}}$.

Утверждение. Функционал Φ дифференцируем по Фреше.

Действительно,

$$\Phi(x+h) = \Phi(x) =$$
$$= \int_a^b f(s, x(s)+h(s), \dot{x}(s)+\dot{h}(s))\, ds - \int_a^b f(s, x(s), \dot{x}(s))\, ds =$$
$$= \int_a^b [f(s, x(s)+h(s), \dot{x}(s)+\dot{h}(s)) - f(s, x(s), \dot{x}(s))]\, ds =$$
$$= \int_a^b [f_x(s,x(s),\dot{x}(s))h(s) + f_{\dot{x}}(s,x(s),\dot{x}(s))\dot{h}(s) + \omega(s,x(s),\dot{x}(s),h(s),\dot{h}(s))]\, ds =$$
$$= \int_a^b [f_x(s,x(s),\dot{x}(s))h(s) + f_{\dot{x}}(s,x(s),\dot{x}(s))\dot{h}(s)]\, ds + \hat{\omega}(x,h) =$$
$$= d\Phi(x,h) + \hat{\omega}(x,h)\,,$$

где

$$d\Phi(x,h) = \int_a^b [f_x(s,x(s),\dot{x}(s))h(s) + f_{\dot{x}}(s,x(s),\dot{x}(s))\dot{h}(s)]\, ds -$$

непрерывный линейный относительно h функционал, дифференциал Фреше, а

$$\hat{\omega}(x,h) = \int_a^b \omega(s, x(s), \dot{x}(s), h(s), \dot{h}(s))\, ds -$$

поправочное слагаемое,

$$|\hat{\omega}(x,h)| \le L(b-a)\|h\|^2_{C^1[a,b]} = o(\|h\|_{C^1[a,b]})\,.$$

(Замечание. Липшицевость является несколько избыточным требованием, но оно удобно для получения оценок и почти всегда выполняется.)

Теперь мы переходим к рассмотрению задачи на поиск экстремума функционала Φ либо на всём пространстве (задача безусловной оптимизации), либо на некотором множестве M (задача условной оптимизации). В этом случае необходимым условием экстремума является обращение в нуль дифференциала Фреше либо на всех функциях $h \in C^1[a,b]$, либо на допустимых, определяемых конкретным видом множества M (в последнем случае речь идёт о внутренних точках множества в оговоренном выше смысле, в противном случае нужно рассматривать неравенства для вариаций).

Далее мы рассмотрим две задачи классического вариационного исчисления: задачу с закремлёнными концами и задачу со свободными концами. Либо одну, либо другую обычно называют "простейшей задачей вариационного исчисления но разные авторы делают разный выбор (одна проще по постановке, другая с точки зрения решения), поэтому мы не будем пользоваться этой формулировкой.

Задача со свободными концами: найти точки экстремума функционала $\Phi(x)$ среди всех $x \in C^1[a,b]$. Это задача на безусловный экстремум, и необходимым условием такого экстремума является выполнение равенства

$$\int_a^b [f_x(s, x(s), \dot{x}(s))h(s) + f_{\dot{x}}(s, x(s), \dot{x}(s))\dot{h}(s)]\, ds = 0$$

при всех $h \in C^1[a,b]$.

Задача с закреплёнными концами: найти точки экстремума функционала $\Phi(x)$ среди всех $x \in C^1[a,b]$, удовлетворяющих краевым условиям

$$x(a) = A\,, \qquad x(b) = B\,,$$

где A и B – заданные числа. Это уже задача условной оптимизации, множество M – множество функций из $C^1[a,b]$, удовлетворяющих краевым условиям. Если хотя бы одно из чисел A или B отлично от нуля, условия неоднородные, и это не подпространство. Гиперплоскость, смещённое подпространство. Выпуклое множество (проверить!).

Ограничения на h: функция $x+h$ должна удовлетворять краевым условиям:

$$x(a) + h(a) = A\,, \qquad x(b) + h(b) = B\,.$$

Поскольку сама функция x должна этим условиям удовлетворять, отсюда следует, что

$$h(a) = h(b) = 0\,.$$

Это условие уже однородно и определяет подпространство $C_0^1[a,b]$ (замкнутое – проверить). Соответственно, интеграл, являющийся дифференциалом Фреше, должен обращаться в нуль на любом элементе этого подпространства $h \in C_0^1[a,b]$.

Для того, чтобы двинуться дальше в рассмотрении этих двух задач, нам понадобится доказать несколько важных вспомогательных утверждений.

Лемма Лагранжа.

Пусть $u \in C[a,b]$, и для любого $h \in C_0^1[a,b]$ справедливо равенство

$$\int_a^b u(s)h(s)\,ds = 0\,.$$

Тогда функция u тождественно равна нулю.

(Замечание. Использованная здесь буква u не имеет отношения к формальному параметру функции f.)

Докажем, что для произвольной функции $u \in C[a,b]$, не равной нулю тождественно, найдётся функция $h \in C_0^1[a,b]$ такая, что интеграл не будет равен нулю.

Действительно, пусть $u(s_0) \neq 0$, где s_0 – некоторая внутренняя точка. Тогда в силу непрерывности u найдётся окрестность этой точки $(\alpha,\beta) \subset (a,b)$, в которой u имеет тот же знак, что и $u(s_0)$. Строим функцию $h \in C_0^1[a,b]$, локализованную на (α,β) и положительную на этом интервале (такие существуют – например, квадратичный сплайн – построить!). Произведение непрерывно и знакопостоянно на (α,β), нуль за пределами интервала, интеграл не равен нулю.

Если s_0, где $u(s_0) \neq 0$ – граничная точка, то по непрерывности найдётся и внутренняя. Лемма доказана.

Лемма Дюбуа-Реймона.

Пусть $v \in C[a,b]$, и для любого $h \in C_0^1[a,b]$ справедливо равенство

$$\int_a^b v(s)\dot{h}(s)\,ds = 0\,.$$

Тогда функция v – константа.

(Замечание. Использованная здесь буква v не имеет отношения к формальному параметру функции f.)

Убедимся сначала, что для константы интеграл равен нулю. Действительно, если $v = C$, то

$$\int_a^b C\dot{h}(s)\,ds = Ch\big|_a^b = 0\,.$$

Теперь приступим к доказательству. Заметим, что если g – некоторая функция на $[a,b]$ то

$$\exists h_g \in C_0^1[a,b] : g(s) = \dot{h}_g(s) \;\Leftrightarrow g \in C[a,b] \wedge \int_a^b g(s)\,ds = 0\,.$$

Действительно, если $g(s) = \dot{h}_g(s)$, то $g \in C[a,b]$ в силу непрерывной дифференцируемости h_g, и

$$\int_a^b g(s)\,ds = \int_a^b \dot{h}_g(s)\,ds = \left. h_g\right|_a^b = 0\,.$$

Обратно, если g – непрерывная функция с нулевым средним, то

$$h_g(s) = \int_a^s g(\tau)\,d\tau \in C_0^1[a,b]\,,$$

поскольку эта функция непрерывно дифференцируема и обращается в нуль на границах.

Теперь обозначим

$$C = \frac{1}{b-a}\int_a^b v(s)\,ds\,-$$

среднее значение функции v на отрезке, и

$$g(s) = v(s) - C\,-$$

непрерывная функция с нулевым средним (проверить!). Тогда для любого $h \in C_0^1[a,b]$, согласно условию леммы,

$$\int_a^b g(s)\dot{h}(s)\,ds = \int_a^b (v(s)-C)\dot{h}(s)\,ds = \int_a^b v(s)\dot{h}(s)\,ds = 0\,.$$

Выберем $h(s) = h_g(s)$, тогда

$$\int_a^b g(s)\dot{h}_g(s)\,ds = \int_a^b g^2(s)\,ds = 0\,,$$

Поскольку подынтегральная функция непрерывна и неотрицательна, отсюда следует, что она тождественно равна нулю на отрезке $[a,b]$, т.е.

$$g^2(s) = (v(s)-C)^2 = 0\,,$$

откуда

$$v(s) = C\,,$$

что и требовалось доказать.

Обобщённая **лемма** Дюбуа-Реймона.

(Замечание. Некоторые авторы под леммой Дюбуа-Реймона понимают именно эту лемму.)

Пусть $u, v \in C[a,b]$, и для любого $h \in C_0^1[a,b]$ справедливо равенство

$$\int_a^b (u(s)h(s) + v(s)\dot{h}(s))\, ds = 0\,.$$

Тогда функция $v \in C^1[a,b]$, и

$$u(s) = \dot{v}(s)\,.$$

Убедимся сначала, что в случае $u(s) = \dot{v}(s)$ интеграл действительно обращается в нуль. Действительно,

$$\int_a^b (u(s)h(s) + v(s)\dot{h}(s))\, ds = \int_a^b (\dot{v}(s)h(s) + v(s)\dot{h}(s))\, ds =$$
$$= \int_a^b (v(s)h(s))^{\cdot}\, ds = \left.(v(s)h(s))\right|_a^b = 0\,.$$

Теперь приступим к доказательству. Обозначим

$$w(s) = \int_a^s u(\tau)\, d\tau\,,$$

тогда $u(s) = \dot{w}(s)$, и $\forall h \in C_0^1[a,b]$:

$$\int_a^b (u(s)h(s) + v(s)\dot{h}(s))\, ds = \int_a^b (\dot{w}(s)h(s) + v(s)\dot{h}(s))\, ds =$$
$$= \left.(w(s)h(s))\right|_a^b + \int_a^b (-w(s)\dot{h}(s) + v(s)\dot{h}(s))\, ds = \int_a^b (-w(s) + v(s))\dot{h}(s)\, ds = 0\,.$$

Тогда, согласно лемме Рюбуа-Реймона,

$$-w(s) + v(s) = C\,,$$

т.е.

$$v(s) = w(s) + C\,.$$

Тогда $v \in C^1[a,b]$, как сумма $w \in C^1[a,b]$ и конствнты. Продифференцировав полученное равенство, получаем:

$$\dot{v}(s) = \dot{w}(s) = u(s)\,,$$

что и требовалось доказать.

Замечание. Если бы было заранее известно, что $v \in C^1[a,b]$, доказательство можно было бы сократить:

$$\int_a^b (u(s)h(s) + v(s)\dot{h}(s))\, ds =$$

$$= (v(s)h(s))|_a^b + \int_a^b (u(s)h(s) - \dot{v}(s)h(s))\, ds =$$

$$= \int_a^b (u(s) - \dot{v}(s))h(s)\, ds = 0\,,$$

и тогда по лемме Лагранжа

$$u(s) - \dot{v}(s) = 0\,.$$

Теперь возвращаемся к рассмотрению задачи с закреплёнными концами. Необходимое условие экстремума в этой задаче в точности повторяет условия из обобщённой леммы Дюбуа-Реймона при

$$u(s) = f_x(s, x(s), \dot{x}(s))\,,$$
$$v(s) = f_{\dot{x}}(s, x(s), \dot{x}(s))\,.$$

Поэтому в силу результата леммы имеем:

$$\frac{d}{ds} f_{\dot{x}}(s, x(s), \dot{x}(s)) = f_x(s, x(s), \dot{x}(s))\,.$$

Принято это уравнение, именуемое уравнением Эйлера (или уравнением Эйлера-Лагранжа) записывать в виде

$$-\frac{d}{ds} f_{\dot{x}}(s, x(s), \dot{x}(s)) + f_x(s, x(s), \dot{x}(s)) = 0\,.$$

Произвольное решение этого уравнения (независимо от того, удовлетворяет ли оно кравевым условиям) называют экстремалью исходного функционала.

Производная $\frac{d}{ds}$ по переменной s – полная, учитывающая зависимость от s всех аргументов:

$$\frac{d}{ds} f_{\dot{x}}(s, x(s), \dot{x}(s)) =$$
$$= f_{s\dot{x}}(s, x(s), \dot{x}(s)) + f_{x\dot{x}}(s, x(s), \dot{x}(s))\dot{x} + f_{\dot{x}\dot{x}}(s, x(s), \dot{x}(s))\ddot{x}$$

(эта формула справедлива, если все производные в правой части существуют). Как мы видим, уравнение Эйлера – обыкновенное дифференциальное уравнение второго порядка относительно $x(s)$, для которой, с учётом условий $x(a) = A\,, x(b) = B$ мы получаем первую краевую задачу, решение которой (в принципе оно может быть не одно) незывается допустимой экстремалью (т.е. экстремалью, допустимой по ограничению) и является стационарной точкой функционала на множестве M и кандидатом на роль условного экстремума. Чтобы выяснить, доставляет ли найденная функция экстремум функционалу, а если да, то какой именно (максимум или минимум), требуется дальнейший анализ, который иногда по трудоёмкости превосходит собственно процесс решения краевой задачи.

Пепейдём к рассмотрению задачи со свободными концами, т.е. задачи безусловной оптимизации, когда на функцию $x(s)$ не накладывается никаких ограничений. В этом случае дифференциал Фреше $d\Phi$ должен обращаться в нуль на любых элементах $h \in C^1[a,b]$, в том числе и на $h \in C_0^1[a,b]$. Отсюда, опять применяя обобщённую лемму Дюбуа-Реймона, мы снова приходим к уравнению Эйлера. Таким образом, любое решение оптимизационной задачи без ограничений является экстремалью функционала Φ.

Для того, чтобы дополнить уравнение Эйлера краевыми условиями, снова выполним интегрирование по частям, теперь уже не считая, что $h \in C_0^1[a,b]$:

$$\begin{gathered}\int_a^b (f_x(s,x(s),\dot{x}(s))h(s) + f_{\dot{x}}(s,x(s),\dot{x}(s))\dot{h}(s))\,ds = \\ = (f_{\dot{x}}(s,x(s),\dot{x}(s))h(s))\big|_a^b + \\ + \int_a^b \left(f_x(s,x(s),\dot{x}(s)) - \frac{d}{ds}f_{\dot{x}}(s,x(s),\dot{x}(s))\right)h(s)\,ds = \\ = f_{\dot{x}}(b,x(b),\dot{x}(b))h(b) - f_{\dot{x}}(a,x(a),\dot{x}(a))h(a) = 0\,.\end{gathered}$$

Последний интеграл исчез, так как подынтегральная функция тождественно равна нулю в силу уравнения Эйлера, осталась подстановка, которая также должна обращаться в нуль при произвольных $h \in C_0^1[a,b]$. Поочерёдно подставляя в последнее равенство функцию $h(s) = s - a$ и $h(s) = s - b$, мы получим, что

$$f_{\dot{x}}(a,x(a),\dot{x}(a)) = f_{\dot{x}}(b,x(b),\dot{x}(b)) = 0\,.$$

Эти условия носят название условий трансверсальности или, в другой терминологии, естественных граничных условий (естественных в том смысле, что мы их не накладываем при постановке задачи, как в задаче с закреплёнными концами, а они выполняются на экстремальном элементе сами собой). Решив краевую задачу для уравнения Эйлера, получим стационарную точку функционала на всём пространстве $C_0^1[a,b]$ (или несколько таких точек). Каждое из решений затем, как и в задаче с закреплёнными концами, нужно исследовать на экстремальность.

Замечание. Условия трансверсальности возникают в качестве краевых условий для уравнения Эйлера и при поиске экстремума функционалов более общего вида, содержащих, наряду с интегральным слагаемым, также слагаемые, зависящие от значений функции x в граничных точках. Для таких функционалов условия трансверсальности имеют более сложный вид.