Лекция 2
16.02.2023

$$\frac{\partial}{\partial x}(x^T A x) = 2Ax.$$

~~Легко проверить это путем перемножения матрицы $A=A^T$ на вектор x, что выражение для производной совпадает с вычисленным путем покомпонентного дифференцирования квадратичной формы.~~

Лекция 2

1.3.2. Пример: стационарные точки заданной функции суммой линейной и квадр. форм

Рассмотрим ~~пример~~

Задана функция n-переменных

$$f(x) = \frac{1}{2} x^T Q x + x^T a, \quad x \in R^n, \ a \in R^n, \ Q = Q^T -$$

ичная матрица.

Пусть $Q > 0$ положительно определенная (т.е $\forall x \neq 0 \quad x^T Q x > 0$), поэтому матрица Q имеет обратную Q^{-1}.

Требуется найти стационарную точку $\hat{x}$ для этой функции т.е такую, что $f'(\hat{x}) = 0$.

Решение. Вычислим производную функции $f(x)$

$$f'(x) = \frac{\partial f(x)}{\partial x} = \frac{\partial}{\partial x}\left(\frac{1}{2} x^T Q x + x^T a\right) =$$

$$= \frac{1}{2} \cdot 2Qx + a = Qx + a$$

Стационарная точка определяется из алгебраического уравнения

$$Q\hat{x} + a = 0,$$

Это уравнение имеет решение так как Q имеет обратную матрицу

$$\hat{x} = -Q^{-1} a.$$

Рассмотрим частный случай этой задачи $n=2$, $Q=I_2$

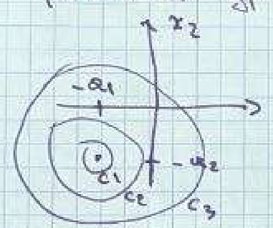
Тогда

$$f(x)=\frac{1}{2}(x_1,x_2)\begin{pmatrix}x_1\\x_2\end{pmatrix}+(x_1\,x_2)\begin{pmatrix}a_1\\a_2\end{pmatrix}=\frac{1}{2}(x_1^2+x_2^2)+a_1x_1+a_2x_2.$$

Стационарная точка

$$\hat{x}=-I_2a=-\begin{pmatrix}a_1\\a_2\end{pmatrix}.$$

На плоскости (x_1,x_2) изобразим линии равного уровня функции $f(x)=c$.



$c_1<c_2<c_3$

Вычислим значение функции в ~~точке~~ стационарной точке

$$f(\hat{x})=\frac{1}{2}\hat{x}^TQ\hat{x}+\hat{x}^Ta=$$

$$=\frac{1}{2}a^TQ^{-1}QQ^{-1}a-a^TQ^{-1}a=$$

$$=-\frac{1}{2}a^TQ^{-1}a=-\frac{1}{2}a^TI_2a=-\frac{1}{2}|a|^2$$

$$=-\frac{1}{2}(a_1^2+a_2^2)$$

Можно найти точку экстремума другим ~~способом~~ непосредственно — путем выделения полного квадрата. Для $n=2$ $Q=I$ получаем

$$f(x)=\frac{1}{2}x^Tx+x^Ta=\frac{1}{2}x^Tx+x^Ta+\frac{1}{2}a^Ta-\frac{1}{2}a^Ta=$$

$$=\frac{1}{2}(x^Tx+x^Ta+a^Tx+a^Ta)-\frac{1}{2}a^Ta=$$

$$=\frac{1}{2}(x+a)^T(x+a)-\frac{1}{2}a^Ta=\frac{1}{2}|x+a|^2-\frac{1}{2}\overset{|a|^2}{a^Ta}$$

Минимум функции достигается когда $\hat{x}=-a$ значение функции в точке экстремума

$$f(\hat{x})=-\frac{1}{2}|a|^2.$$

2. Бесконечномерные экстремальные задачи. Простейшая задача вариационного исчисления.

(АТФ) § 1.4 стр. 54

Вскоре после работы И. Бернулли о брахистохроне (1696 г) стали появляться и решаться многие задачи подобного типа. И. Бернулли поставил перед своим учеником Л. Эйлером проблему найти общий способ решения таких задач. В 1744 г. вышел труд Л. Эйлера в котором были заложены основы нового раздела математического анализа — вариационное исчисление. Аппроксимируя кривые ломаными и переходя к пределу, Эйлер вывел дифференциальное уравнение второго порядка которому должны удовлетворять экстремали — линии подозрительные на экстремум. В последствии Лагранж назвал его уравнением Эйлера.

В 1759 году появилась первая работа Лагранжа и с ней новые методы исследования экстремальных задач. Лагранж варьирует кривую подозреваемую на экстремум, выделяет из приращения главную линейную часть, которую называют вариацией, и пользуется тем, что в точке экстремума вариация должна обращаться в ноль.

Метод Лагранжа становится в последствии общепринятым. Этим методом мы в последствии выведем уравнение Эйлера. После работ Лагранжа по предложению Эйлера весь раздел математики в котором применяется метод Лагранжа стали называть вариационным исчислением.

2.1. Простейшая задача классического вариационного исчисления.

Определение. (функционала)

Отображение

$$J(x(\cdot)): C^1_{[t_0,t_1]} \to R^1$$

гладкой непрерывно дифференцируемой функции на числовую ось называют функционалом.

В классической простейшей задаче вариационного исчисления ~~применяется~~ используют функционал вида

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

где $x(\cdot)$ – аргумент функционала, непрерывно дифференцируемая функция, $L(t, x, \dot{x})$ функция трех переменных.

~~Задача классического вариационного исчисления имеет следующую формулировку~~

Определение: Непрерывно дифференцируемые функции $x(t) \in C^1_{[t_0, t_1]}$, которая для фиксированных t_0 начального и t_1 конечного моментов времени принимает фиксированные значения $x(t_0) = x_0$ и $x(t_1) = x_1$ будем называть допустимыми.

Определение: Функция $\hat{x}(\cdot)$ допустимая функция $\hat{x}(\cdot)$ доставляет локальный минимум (максимум) функционалу (2.1), если найдётся такое число $\varepsilon > 0$, что для любой функции $x(\cdot) \in C^1_{[t_0, t_1]}$ ~~для~~ которой $x(t_0) = 0$, $x(t_1) = 0$ $\| x(\cdot) \| < \varepsilon$ выполнено неравенство

$$J(\hat{x}(\cdot) + x(\cdot)) \geqslant J(\hat{x}(\cdot)) \quad (\leqslant J(x(\cdot))).$$

Задача классического вариационного исчисления (простейшая задача вариационного исчисления).

Определить такую допустимую функцию $\hat{x}(\cdot)$ которая доставляет ~~экстремум функционалу~~ локальный минимум (максимум) функционалу (2.1)

Теорема: (Необходимое условие экстремума в простейшей задаче вариационного исчисления).

Пусть $\hat{x}(\cdot)$ допустимая функция доставляет минимум (максимум) функционалу

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

Тогда (=>) она удовлетворяет уравнению

$$(2.2) \quad -\frac{d}{dt} L_{\dot{x}}(t, \hat{x}(t), \dot{\hat{x}}(t)) + L_x(t, \hat{x}(t), \dot{\hat{x}}(t)) = 0.$$

Уравнение (2.2) называется уравнением Эйлера.

<u>Определение</u>. Допустимая функция $\hat{x}(\cdot)$ для которой выполнено уравнение Эйлера называется стационарной (простейшей задачи вариационного исчисления) точкой или экстремалью.

<u>Следствие:</u> Из теоремы следует, что точка локального экстремума $\hat{x}(\cdot)$ является экстремалью. Обратное неверно.

<u>Доказательство теоремы</u>.

Будем производить в соответствии с классическим способом предложенным Лагранжем. В этом случае требуется ввести дополнительное предположение о том, что функция $L_{\dot{x}}(t) = L_{\dot{x}}(t, x(t), \dot{x}(t))$ непрерывно дифференцируема по t.

Доказательство теоремы состоит из трех частей этапов.

и она определяется как

$$\delta J(\hat{x}(\cdot);x(\cdot)) = \lim_{\lambda\to 0}\left(\frac{J(\hat{x}(\cdot)+\lambda x(\cdot)) - J(\hat{x}(\cdot))}{\lambda} - \frac{r(\lambda x(\cdot))}{\lambda}\right) =$$

А: Этап I. Определение первой вариации функционала

Проварьируем функцию $\hat{x}(\cdot)$ определяющую экстремум функционала $J(\hat{x}(\cdot))$, задав отклонение от $\hat{x}(\cdot)$ в виде $\hat{x}(\cdot)+\lambda x(\cdot)$, где $\lambda \in R^1$ варьируемая переменная, $x(\cdot)$ – произвольная фиксированная непрерывно дифференцируемая функция ($x(\cdot) \in C^1_{[t_0,t_1]}$). Значение функционала вычисленное для этой функции обозначим как $J(\hat{x}(\cdot)+\lambda x(\cdot))$.

Определение: (Вариация функционала) Функционал называется дифференцируем в точке $\hat{x}(\cdot)$ если справедливо

$$J(\hat{x}(\cdot)+\lambda x(\cdot)) = J(\hat{x}(\cdot)) + \delta J(\hat{x}(\cdot);x(\cdot))\cdot\lambda + r(\lambda x(\cdot))$$

где $r(\lambda x(\cdot)) = o(\lambda)$

для любой $x(\cdot) \in C^1_{[t_0,t_1]}$.

а $\delta J(\hat{x}(\cdot);x(\cdot))$ линейный функционал по $x(\cdot)$ (при фиксированом $\hat{x}(\cdot)$) его называют вариацией функционала в точке $\hat{x}(\cdot)$.

~~Он вычисляется~~ Вариация функционала это производная Фреше функционала, что является аналогом ~~частной~~ производной ~~для~~ функции нескольким переменные для бесконечномерного случая, ~~т.е.~~ функционала.

и она определяется как

$$\delta J(\hat{x}(\cdot); x(\cdot)) = \lim_{\lambda \to 0} \left(\frac{J(\hat{x}(\cdot)+\lambda x(\cdot)) - J(\hat{x}(\cdot))}{\lambda} - \frac{r(\lambda x(\cdot))}{\lambda} \right) =$$

$$= \lim_{\lambda \to 0} \frac{J(\hat{x}(\cdot)+\lambda x(\cdot)) - J(\hat{x}(\cdot))}{\lambda}$$

Заметим, что

$J(\hat{x}(\cdot)+\lambda x(\cdot)) = \varphi(\lambda; \hat{x}(\cdot), x(\cdot))$ функция одной переменной λ, ~~так как $\hat{x}(\cdot)$ и $x(\cdot)$ фиксированные функции~~

$$\varphi(\lambda; \hat{x}(\cdot), x(\cdot)) : \lambda \cancel{x(\cdot)} \to R^1$$

при фиксированных $\hat{x}(\cdot)$ и $x(\cdot)$

и поэтому

$$\delta J(\hat{x}(\cdot); x(\cdot)) = \lim_{\lambda \to 0} \frac{\varphi(\lambda; \hat{x}(\cdot), x(\cdot)) - \varphi(0; \hat{x}(\cdot), x(\cdot))}{\lambda}$$

$$= \varphi'(0; \hat{x}(\cdot), x(\cdot)),$$

~~и её можно~~ вариация функционала в точке $\hat{x}(\cdot)$ совпадает с первой производной функции $\varphi(\lambda; \hat{x}(\cdot), x(\cdot))$ вычисленной для $\lambda = 0$ при фиксированных $\hat{x}(\cdot), x(\cdot)$

Вычислим производную функции для функционала заданного в условии теоремы

$$\varphi(\lambda; \hat{x}(\cdot), x(\cdot)) = J(\hat{x}(\cdot)+\lambda x(\cdot)) =$$

$$= \int_{t_0}^{t_1} L(t, \hat{x}(t)+\lambda x(t), \dot{\hat{x}}(t)+\lambda \dot{x}(t))\, dt$$

При вычислении производной под знаком интеграла по λ ~~даже~~ предполагается, что все необходимые частные производные непрерывны.

$$\varphi'(\lambda; \hat{x}(\cdot), x(\cdot)) = \int_{t_0}^{t_1} \Big(\frac{\partial L(t,\hat{x},\dot{\hat{x}})}{\partial t} \cdot \frac{\partial t}{\partial \lambda} + \frac{\partial L(t,\hat{x},\dot{\hat{x}})}{\partial x} \cdot \frac{\partial(\hat{x}+\lambda x)}{\partial \lambda} + \frac{\partial L(t,\hat{x},\dot{\hat{x}})}{\partial \dot{x}} \cdot \frac{\partial(\dot{\hat{x}}+\lambda \dot{x})}{\partial \lambda} \Big) dt$$

Введем обозначения

$$\frac{\partial L(t,\hat{x}(t),\dot{\hat{x}}(t))}{\partial x} = q(t),$$

$$\frac{\partial L(t,\hat{x}(t),\dot{\hat{x}}(t))}{\partial \dot{x}} = p(t).$$

Вычислим частные производные

$$\frac{\partial t}{\partial \lambda} = 0; \quad \frac{\partial(\hat{x}+\lambda x)}{\partial \lambda} = x(t); \quad \frac{\partial(\dot{\hat{x}}+\lambda \dot{x})}{\partial \lambda} = \dot{x}(t).$$

Тогда $\displaystyle \varphi'(\lambda; \hat{x}(\cdot), x(\cdot)) = \int_{t_0}^{t_1} (q(t)x(t) + p(t)\dot{x}(t))dt$

В результате мы получили выражение для вариации функционала в точке $\hat{x}(\cdot)$ для произвольного $x(\cdot)$

~~$\delta J(\hat{x}(\cdot); x(\cdot)) = \varphi'(\lambda; \hat{x}(\cdot), x(\cdot))$~~

она не зависит от λ т.е

$$\varphi'(\lambda; \hat{x}(\cdot), x(\cdot)) = \varphi'(0; \hat{x}(\cdot), x(\cdot))$$

и является линейным функционалом от $x(\cdot)$, поэтому вариация функционала определяется как

$$\delta J(\hat{x}(\cdot); x(\cdot)) = \varphi'(0; \hat{x}(\cdot), x(\cdot)) = \int_{t_0}^{t_1} (q(t)x(t) + p(t)\dot{x}(t))dt$$

где $\quad q(t) = L_x(t, \hat{x}(t), \dot{\hat{x}}(t))$

$\quad\quad p(t) = L_{\dot{x}}(t, \hat{x}(t), \dot{\hat{x}}(t))$

Б. Этап II. Преобразование первой вариации с помощью интегрирования по частям

допустимо заметим, что $\hat{x}(\cdot)$ - допустимая ф-я, для того чтобы $\hat{x}(\cdot)+\lambda x(\cdot)$ - была

Пусть требуется, чтобы $x(t_0)=x(t_1)=0$, функция $p(t)$ непрерывно дифференцируемая.

Проинтегрируем по частям второй член в выражении $\delta J(\hat{x}(\cdot); x(\cdot))$.

Воспользуемся формулой интегрирования по частям

$$\int_{t_0}^{t_1}\frac{d(u(t)\cdot v(t))}{dt}\cdot dt=\int_{t_0}^{t_1}\left(\frac{du}{dt}\cdot v(t)+u(t)\frac{dv(t)}{dt}\right)dt$$

Для того, чтобы воспользоваться формулой введем следующие обозначения

$$p(t)=u(t),\quad \dot{x}(t)=\frac{dv(t)}{dt},\quad \frac{dp(t)}{dt}=\frac{du(t)}{dt};\quad x(t)=v(t)$$

Вычислим второй член выражения вариации $\delta J(\hat{x}(\cdot); x(\cdot))$, по формуле интегрирования по частям

$$\int_{t_0}^{t_1}\frac{d(p(t)\cdot x(t))}{dt}\cdot dt=\int_{t_0}^{t_1}\left(\frac{dp(t)}{dt}x(t)+p(t)\frac{dx(t)}{dt}\right)dt$$

отсюда имеем

$$\int_{t_0}^{t_1}p(t)\,\dot{x}(t)\,dt=\int_{t_0}^{t_1}d(p(t)x(t))-\int_{t_0}^{t_1}\dot{p}(t)x(t)\,dt$$

~~где~~ интеграл от полного дифференциала легко вычисляется, в результате получаем

$$\int_{t_0}^{t_1}p(t)\,\dot{x}(t)\,dt=p(t)\,x(t)\Big|_{t=t_0}^{t_1}-\int_{t_0}^{t_1}\dot{p}(t)\,x(t)\,dt$$

Учитывая, что $x(t_0) = x(t_0) = 0$ получаем

$$\int_{t_0}^{t_1} p(t)\dot{x}(t)dt = -\int_{t_0}^{t_1} \dot{p}(t)x(t)dt.$$

Подставляя полученный результат в выражение для первой вариации функционала получаем

$$\delta J(\hat{x}(\cdot); x(\cdot)) = \int_{t_0}^{t_1} (-\dot{p}(t) + q(t))x(t)dt = \int_{t_0}^{t_1} a(t)x(t)dt$$

где $a(t) = -\dot{p}(t) + q(t)$. А $\delta J(\hat{x}(\cdot); x(\cdot))$ линейный функционал.

По условию теоремы $\hat{x}(\cdot)$ доставляет локальный экстремум (пусть это минимум) в задаче (1), поэтому функция одной переменной λ

$$J(\hat{x}(\cdot) + \lambda x(\cdot)) = \varphi(\lambda; \hat{x}(\cdot), x(\cdot))$$

имеет локальный минимум при $\lambda = 0$.

По теореме Ферма в точке локального минимума $\varphi'(0) = 0$, поэтому

$$\varphi'(0) = \delta J(\hat{x}(\cdot); x(\cdot)) = 0$$

для любого $x(t) \in C^1_{[t_0, t_1]}$, если $x(t_0) = x(t_1) = 0$.

С учетом полученного выражения для первой вариации функционала по т. Ферма имеем

$$\varphi'(0) = \delta J(\hat{x}(\cdot); x(\cdot)) = \int_{t_0}^{t_1} a(t)x(t)dt = 0$$

получаем для точки экстремума $\hat{x}(\cdot)$ (для экстремали)

$$\int_{t_0}^{t_1} a(t)x(t)dt = 0.$$

С. Шаг 3. Основная лемма классического вариационного исчисления (Лемма Лагранжа)

Лемма (Лагранжа)

Пусть непрерывная функция $a(t) \in C_{[t_0,t_1]}$ обладает свойством

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

для любой непрерывно дифференцируемой функции $x(\cdot) \in C^1_{[t_0,t_1]}$ у которой $x(t_0) = x(t_1) = 0$.

Тогда ($\Rightarrow$)

$$a(t) \equiv 0.$$

Доказательство.

Пусть $a(\tau) \neq 0$ в некоторой точке $\tau \in [t_0, t_1]$. Тогда в силу непрерывности $a(\cdot)$ найдется отрезок $\Delta = [\tau_1, \tau_2] \subset [t_0, t_1]$ на котором $a(t) \neq 0$, $\tau \in \Delta$. Пусть $a(t) \geq m > 0$, $\tau \in \Delta$.

Построим функцию

$$\tilde{x}(t) = \begin{cases} (t-\tau_1)^2 (t-\tau_2)^2, & t \in \Delta \\ 0, & t \notin \Delta \end{cases}$$

Легко проверить, что $\tilde{x}(t) \in C^1_{[t_0,t_1]}$ и $\tilde{x}(t_0) = \tilde{x}(t_1) = 0$. По условию леммы

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

С другой стороны

$$\int_{t_0}^{t_1} a(t)\, \tilde{x}(t)\, dt = \int_{\tau_1}^{\tau_2} a(t)\, \tilde{x}(t)\, dt \geq \int_{\tau_1}^{\tau_2} m\, \tilde{x}(t)\, dt > 0$$

Это противоречие доказывает лемму.

Применим лемму к полученному условию стационарности для первой вариации функционала

$$\int_{t_0}^{t_1} a(t)x(t)dt = 0 \quad , \; x \in C^1_{[t_0,t_1]} \; x(t_0) = x(t_1) = 0,$$

и получаем

$$a(t) = -\dot{p}(t) + q(t) \equiv 0.$$

или окончательно, получим утверждение теоремы

$$-\frac{d}{dt} L_{\dot{x}}(t, \hat{x}(t), \dot{\hat{x}}(t)) + L_x(t, \hat{x}(t), \dot{\hat{x}}(t)) = 0$$

2.2. Векторная задача

Выше была рассмотрена одномерная простейшая задача вариационного исчисления. Аналогично ставится и решается векторная задача.

Функционал в векторном случае имеет вид

$$J(x_1(\cdot), \dots, x_n(\cdot)) = \int_{t_0}^{t_1} L(t, x_1(\cdot), \dots, x_n(\cdot), \dot{x}_1(\cdot), \dots, \dot{x}_n(\cdot)) dt,$$

Если введём вектор (строку)

$$x(t) = (x_1(t), \dots, x_n(t))$$

то функционал запишется компактно

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

т.е. по форме он ничем не отличается от одномерной задачи, надо только понимать, что $x(t)$ и $\dot{x}(t)$ n-мерные вектора элементами которых являются функции, а $L(t,x(\cdot),\dot{x}(\cdot))$ функция $2n+1$ переменных.

Необходимое условие экстремума здесь имеет вид

$$-\frac{d}{dt}L_{\dot{x}_i}(t,\hat{x}_1(t),\dots,\hat{x}_n(t),\dot{\hat{x}}_1(t),\dots,\dot{\hat{x}}_n(t)) +$$
$$+ L_{x_i}(t,\hat{x}_1(t),\dots,\hat{x}_n(t),\dot{\hat{x}}_1(t),\dots,\dot{\hat{x}}_n(t)) = 0$$
$$i=1,\dots,n$$

Эти n уравнений можно записать в векторной форме

$$-\frac{d}{dt}L_{\dot{x}}(t,\hat{x}(t),\dot{\hat{x}}(t)) + L_x(t,\hat{x}(t),\dot{\hat{x}}(t)) = 0,\ x\in R^n$$

здесь $L_{\dot{x}}\in R^n$, $L_x\in R^n$ соответствующие векторные функции векторных аргументов. Т.е. уравнения Эйлера имеют тот же самый вид, но их надо понимать векторно.

Пример. Задача о брахистохроне