Rachunek wariacyjny

Bartosz Wróblewski

26.10.13

Rachunek wariacyjny - dziedzina analizy matematycznej zajmująca się znajdowaniem ekstremów i wartości stacjonarnych funkcjonałów.

Powstał jako odpowiedź na pewne szczególne rozważania w mechanice teoretycznej. Swą nazwę zawdzięcza pierwszej poważnej pracy na ten temat, *Elementa Calculi Variationum* napisanej przez Eulera w 1733 roku. W późniejszym okresie rozwinął się w obszerną naukę stanowiącą podstawę znacznej części fizyki teoretycznej.

Funkcjonał

Funkcjonałem będziemy nazywać funkcję $I: \ D \mapsto$ gdzie D jest zbiorem funkcji klasy C^1 określonych na odcinku domkniętym $[a, b]$. Zbiór takich funkcji będziemy traktować jako przestrzeń metryczną z metryką zadaną przez normę:

$$||y||_0 = \max_{a \leqslant x \leqslant b} |y(x)|$$

albo:

$$||y||_1 = \max_{a \leqslant x \leqslant b} |y(x)| + \max_{a \leqslant x \leqslant b} |y'(x)|$$

w zależności od potrzeb.

Funkcjonał nazywamy **liniowym** jeżeli jest on ciągły oraz dla dowolnych$y_1, y_2 \in D$ $\alpha, \beta \in$ zachodzi:

$$I[\alpha y_1 + \beta y_2] = \alpha I[y_1] + \beta I[y_2]$$

Przyrostem funkcjonału dla ustalonej funkcji y nazwiemy:

$$\Delta I[h] = I[y+h] - I[y]$$

odpowiadający przyrostowi h zmiennej niezależnej (dla tego funkcjonału) y. Jest to funkcjonał zależny od h, w ogólnym przypadku nieliniowy.

Przyrostem funkcjonału dla ustalonej funkcji y nazwiemy:

$$\Delta I[h] = I[y+h] - I[y]$$

odpowiadający przyrostowi h zmiennej niezależnej (dla tego funkcjonału) y. Jest to funkcjonał zależny od h, w ogólnym przypadku nieliniowy.

Różniczka funkcjonału

Różniczką lub wariacją δI funkcjonału I nazywamy główną liniową cześć przyrostu ΔI, czyli funkcjonał liniowy $\varphi(h)$ spełniający:

$$\Delta I[h] = \varphi[h] + \alpha[h]||h||$$

gdzie $\alpha \to 0$ gdy $h \to 0$.

Zadanie 1

Udowodnić, że różniczka funkcjonału (gdy istnieje) jest wyznaczona jednoznacznie.

Zadanie 1

Udowodnić, że różniczka funkcjonału (gdy istnieje) jest wyznaczona jednoznacznie.

Funkcjonał $I[y]$ osiąga **ekstremum** dla $y = y_0$ jeżeli istnieje $\varepsilon > 0$ taki, że dla każdego $h: \; ||h|| < \varepsilon$ wyrażenie $I[y_0 + h] - I[y_0]$ zachowuje znak.
Gdy używamy normy $|| \cdot ||_0$ ekstremum takie nazywamy *silnym*, w przypadku normy $|| \cdot ||_1$ mamy do czynienia z ekstremum *słabym*.

Zadanie 1

Udowodnić, że różniczka funkcjonału (gdy istnieje) jest wyznaczona jednoznacznie.

Funkcjonał $I[y]$ osiąga **ekstremum** dla $y = y_0$ jeżeli istnieje $\varepsilon > 0$ taki, że dla każdego $h : \; ||h|| < \varepsilon$ wyrażenie $I[y_0 + h] - I[y_0]$ zachowuje znak.
Gdy używamy normy $|| \cdot ||_0$ ekstremum takie nazywamy *silnym*, w przypadku normy $|| \cdot ||_1$ mamy do czynienia z ekstremum *słabym*.

Twierdzenie

Warunkiem koniecznym na to, aby funkcjonał $I[h]$ osiągał dla $y = y_0$ ekstremum, jest, aby różniczka (o ile istnieje) była równa zeru.

Funkcję y_0 dla której $\delta I \equiv 0$ nazywamy **ekstremalą**.

Załóżmy, że funkcjonał $I[y]$ osiąga ekstremum dla y_0 oraz $\delta I \not\equiv 0$. Nie zmienia się zatem znak wyrażenia:

$$\Delta I = I[y_0 + h] - I[h]$$

Załóżmy, że funkcjonał $I[y]$ osiąga ekstremum dla y_0 oraz $\delta I \not\equiv 0$. Nie zmienia się zatem znak wyrażenia:

$$\Delta I = I[y_0 + h] - I[h]$$

Jednakże, z założenia istnienia różniczki

$$I[y_0 + h] - I[h] = \delta I[h] + \alpha[h] \cdot ||h||$$

Składnik $\alpha[h] \cdot ||h||$ jest wartością nieskończenie małą niższego rzędu niż $\delta I[h]$ gdy $||h|| \to 0$. Zatem dla dostatecznie małych $||h||$ to różniczka decyduje o znaku wyrażenia.

Załóżmy, że funkcjonał $I[y]$ osiąga ekstremum dla y_0 oraz $\delta I \not\equiv 0$. Nie zmienia się zatem znak wyrażenia:

$$\Delta I = I[y_0 + h] - I[h]$$

Jednakże, z założenia istnienia różniczki

$$I[y_0 + h] - I[h] = \delta I[h] + \alpha[h] \cdot ||h||$$

Składnik $\alpha[h] \cdot ||h||$ jest wartością nieskończenie małą niższego rzędu niż $\delta I[h]$ gdy $||h|| \to 0$. Zatem dla dostatecznie małych $||h||$ to różniczka decyduje o znaku wyrażenia.

Z liniowości otrzymujemy natomiast:

$$\delta I[-h] = -\delta I[h]$$

A więc w dowolnym otoczeniu ΔI może być zarówno dodatnie jak i ujemne. Sprzeczność.

Podstawowy lemat rachunku wariacyjnego

Jeżeli $f(x)$ jest funkcją ciągłą na odcinku $[a, b]$ która dla każdej funkcji ciągłej $h(x)$ spełnia równanie:

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

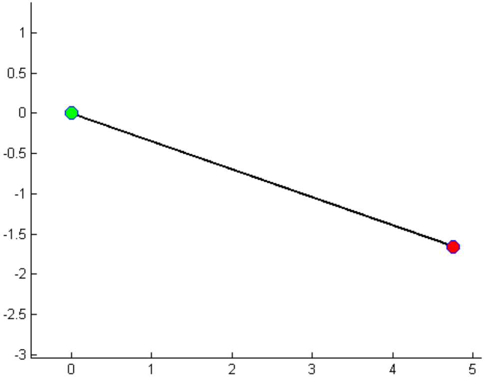
to wtedy: $f(x) \equiv 0$

Krzywa najkrótszego spadku

Dla danych dwóch punktów A, B w jednorodnym polu grawitacyjnym znaleźć taką krzywą, którą tocząca się bez tarcia kulka pokona w najkrótszym czasie. Krzywą taką nazywamy brachistochroną.

Krzywa najkrótszego spadku

Dla danych dwóch punktów A, B w jednorodnym polu grawitacyjnym znaleźć taką krzywą, którą tocząca się bez tarcia kulka pokona w najkrótszym czasie. Krzywą taką nazywamy brachistochroną.



Wprowadźmy dogodny dla nas układ współrzędnych: oś Oy skierowana do dołu, $A = (0,0)$ $B = (x_0, y_0)$. Wyprowadźmy wzór na czas ruchu po krzywej:

$$t = \int t = \int \frac{s}{v} = \int \frac{\sqrt{1 + (y')^2}}{v} x$$

Wprowadźmy dogodny dla nas układ współrzędnych: oś Oy skierowana do dołu, $A = (0,0)$ $B = (x_0, y_0)$. Wyprowadźmy wzór na czas ruchu po krzywej:

$$t = \int t = \int \frac{s}{v} = \int \frac{\sqrt{1+(y')^2}}{v} x$$

Rozważany system jest systemem izolowanym, a jednorodne pole grawitacyjne jest polem o potencjale mgy zatem możemy skorzystać z zasady zachowania energii mechanicznej:

$$\frac{mv^2}{2} = mgy \Rightarrow v = \sqrt{2gy}$$

Podstawiając do poprzedniego równania i ustalając przedziały całkowania otrzymujemy wzór:

$$t = \frac{1}{\sqrt{2g}} \int_0^{x_0} \sqrt{\frac{1+(y')^2}{y}} x$$

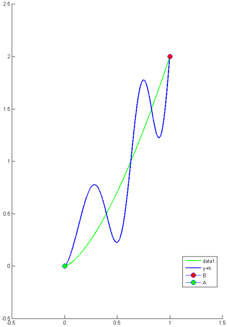
Otrzymany wzór przypisuje każdej krzywej łączącej punkty A, B wartość liczbową odpowiadającą czasie spadku. Definiuje zatem funkcjonał w najprostszej postaci rozważanej w rachunku wariacyjnym:

$$I[y] = \int_a^b F(x, y, y')x$$

Otrzymany wzór przypisuje każdej krzywej łączącej punkty A, B wartość liczbową odpowiadającą czasie spadku. Definiuje zatem funkcjonał w najprostszej postaci rozważanej w rachunku wariacyjnym:

$$I[y] = \int_a^b F(x, y, y')x$$

Wyznaczymy teraz ogólną metodę znajdowania ekstremal funkcjonałów tego typu. Niech h będzie funkcją klasy C^1 taką, że $h(a) = h(b) = 0$. Funkcję taką często nazywa się wariacją funkcji.



Liczymy przyrost funkcjonału:

$$\Delta I[h] = I[y+h] - I[y] = \int_a^b F(x,y,y') - F(x,y+h,y'+h')x =$$
$$= \int_a^b \frac{\partial F}{\partial y}(x,y,y')h + \frac{\partial F}{\partial y'}(x,y,y')h' + o\left(\sqrt{h^2 + h'^2}\right))x$$

Liczymy przyrost funkcjonału:

$$\Delta I[h] = I[y+h] - I[y] = \int_a^b F(x,y,y') - F(x,y+h,y'+h')x =$$
$$= \int_a^b \frac{\partial F}{\partial y}(x,y,y')h + \frac{\partial F}{\partial y'}(x,y,y')h' + o\left(\sqrt{h^2+h'^2}\right))x$$

Z jednoznaczności różniczki funkcjonału otrzymujemy:

$$\delta I = \int_a^b \frac{\partial F}{\partial y}(x,y,y')h + \frac{\partial F}{\partial y'}(x,y,y')h'\,x$$

Liczymy przyrost funkcjonału:

$$\Delta I[h] = I[y+h] - I[y] = \int_a^b F(x,y,y') - F(x,y+h,y'+h')x =$$
$$= \int_a^b \frac{\partial F}{\partial y}(x,y,y')h + \frac{\partial F}{\partial y'}(x,y,y')h' + o\left(\sqrt{h^2+h'^2}\right))x$$

Z jednoznaczności różniczki funkcjonału otrzymujemy:

$$\delta I = \int_a^b \frac{\partial F}{\partial y}(x,y,y')h + \frac{\partial F}{\partial y'}(x,y,y')h'\, x$$

Drugi składnik całkujemy przez części:

$$\int_a^b \frac{\partial F}{\partial y'}(x,y,y')h'\, x = \left.\frac{\partial F}{\partial y'}h\right|_a^b - \int_a^b \frac{}{x}\left(\frac{\partial F}{\partial y'}(x,y,y')\right) h\, x$$
$$= -\int_a^b \frac{}{x}\left(\frac{\partial F}{\partial y'}(x,y,y')\right) h\, x$$

$$\delta I = \int_a^b \left(\frac{\partial F}{\partial y}(x, y, y') + \frac{\partial F}{\partial y'}(x, y, y') \right) h\, x$$

Aby funkcja y była ekstremalą musi zachodzić $\delta I = 0 \; \forall h$. Zatem na mocy **Lematu** otrzymujemy:

$$\delta I = \int_a^b \left(\frac{\partial F}{\partial y}(x, y, y') + \frac{\partial F}{\partial y'}(x, y, y') \right) h\, x$$

Aby funkcja y była ekstremalą musi zachodzić $\delta I = 0 \, \forall h$. Zatem na mocy **Lematu** otrzymujemy:

Równanie Eulera-Lagrange'a

Warunkiem koniecznym na to, aby funkcjonał:

$$I[y] = \int_a^b F(x, y, y') x$$

określony na zbiorze funkcji $y = y(x)\, a \leqslant x \leqslant b$ klasy C^1 spełniających $y(a) = a\, y(b) = B$ osiągał dla danej funkcji $y(x)$ ekstremum, jest, aby ta funkcja spełniała równanie Eulera-Lagrange'a:

$$\frac{\partial F}{\partial y} - \frac{}{x} \frac{\partial F}{\partial y'} = 0$$

Rozpatrzmy przypadki w których równanie to upraszcza się:

Rozpatrzmy przypadki w których równanie to upraszcza się:

- $F = F(x, y)$ wtedy $\frac{\partial F}{\partial y'} = 0$ a zatem:

$$\frac{\partial F}{\partial y} = 0$$

  - zwykłe równanie uwikłane określające jedną lub kilka krzywych

Rozpatrzmy przypadki w których równanie to upraszcza się:

- $F = F(x, y)$ wtedy $\frac{\partial F}{\partial y'} = 0$ a zatem:

$$\frac{\partial F}{\partial y} = 0$$

  - zwykłe równanie uwikłane określające jedną lub kilka krzywych

- $F = F(x, y')$ wtedy $\frac{}{x}\frac{\partial F}{\partial y'} = 0$ a zatem:

$$\frac{\partial F}{\partial y'} = C$$

  - równanie różniczkowe pierwszego rzędu nie zawierające y, a zatem w postaci: $y' = f(x, C)$

Rozpatrzmy przypadki w których równanie to upraszcza się:

- $F = F(x, y)$ wtedy $\frac{\partial F}{\partial y'} = 0$ a zatem:

$$\frac{\partial F}{\partial y} = 0$$

  - zwykłe równanie uwikłane określające jedną lub kilka krzywych

- $F = F(x, y')$ wtedy $\frac{}{x}\frac{\partial F}{\partial y'} = 0$ a zatem:

$$\frac{\partial F}{\partial y'} = C$$

  - równanie różniczkowe pierwszego rzędu nie zawierające y, a zatem w postaci: $y' = f(x, C)$

- $F = F(y, y')$ w tym przypadku:

$$\frac{\partial F}{\partial y} - \frac{}{x}\frac{\partial F}{\partial y'} = \frac{\partial F}{\partial y} - \frac{\partial^2 F}{\partial^2 y'}y'' - \frac{\partial^2 F}{\partial y' \partial y}y'$$

Zauważmy, że:

$$\frac{}{x}\left(F - y'\frac{\partial F}{\partial y'}\right) =$$
$$= \frac{\partial F}{\partial x} + \frac{\partial F}{\partial y}y' + \frac{\partial F}{\partial y'}y'' - y''\frac{\partial F}{\partial y'} - y'\frac{\partial^2 F}{\partial^2 y}y' - y'\frac{\partial^2 F}{\partial y \partial y'}y'' =$$
$$= y'\left(\frac{\partial F}{\partial y} - \frac{\partial^2 F}{\partial^2 y'}y'' - \frac{\partial^2 F}{\partial y' \partial y}y'\right) = y'\left(\frac{\partial F}{\partial y} - \frac{}{x}\frac{\partial F}{\partial y'}\right) = 0$$

Zauważmy, że:

$$\frac{}{x}\left(F - y'\frac{\partial F}{\partial y'}\right) =$$
$$= \frac{\partial F}{\partial x} + \frac{\partial F}{\partial y}y' + \frac{\partial F}{\partial y'}y'' - y''\frac{\partial F}{\partial y'} - y'\frac{\partial^2 F}{\partial^2 y}y' - y'\frac{\partial^2 F}{\partial y \partial y'}y'' =$$
$$= y'\left(\frac{\partial F}{\partial y} - \frac{\partial^2 F}{\partial^2 y'}y'' - \frac{\partial^2 F}{\partial y' \partial y}y'\right) = y'\left(\frac{\partial F}{\partial y} - \frac{}{x}\frac{\partial F}{\partial y'}\right) = 0$$

A więc otrzymujemy równanie w postaci:

$$F - y'\frac{\partial F}{\partial y'} = C$$

- równanie różniczkowe pierwszego rzędu (*całka pierwsza* rówania E-L) Fakt ten czasami nazywany jest **tożsamością Beltramiego.**

Wróćmy teraz do rozważań na temat brachistochrony. Zauważmy że możemy skorzystać z tożsamości Beltramiego. Otrzymujemy zatem:

$$\sqrt{\frac{1+(y')^2}{y}} - \frac{y' \cdot 2y'}{2\sqrt{(1+(y')^2)\,y}} = C$$

Wróćmy teraz do rozważań na temat brachistochrony. Zauważmy że możemy skorzystać z tożsamości Beltramiego. Otrzymujemy zatem:

$$\sqrt{\frac{1+(y')^2}{y}} - \frac{y' \cdot 2y'}{2\sqrt{(1+(y')^2)\,y}} = C$$

Zatem po przekształceniach:

$$y\left(1+(y')^2\right) = C'$$

Wróćmy teraz do rozważań na temat brachistochrony. Zauważmy że możemy skorzystać z tożsamości Beltramiego. Otrzymujemy zatem:

$$\sqrt{\frac{1+(y')^2}{y}} - \frac{y' \cdot 2y'}{2\sqrt{(1+(y')^2)\, y}} = C$$

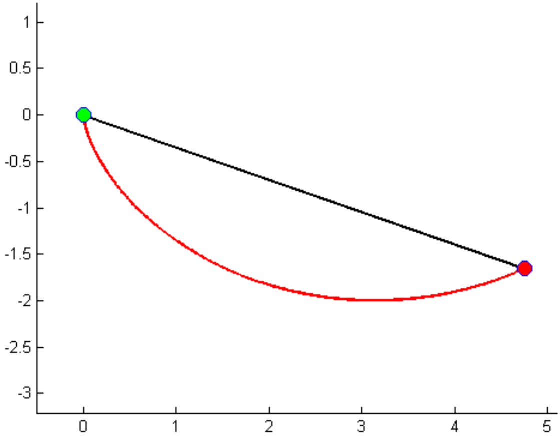
Zatem po przekształceniach:

$$y\left(1+(y')^2\right) = C'$$

Zadanie 2

Rozwiązać wyprowadzone równanie różniczkowe wprowadzając parametr t i przyjmując $y' = \operatorname{ctg} t$. Dowieść że rozwiązaniem tym jest fragment cykloidy.

Rysunek : Porównanie prostej i brachistochrony łączących dwa punkty



Ciekawostka: Brachistchrona jest także *izochroną*.

Uogólnienie na przypadek wielowymiarowy

Analizując dowód łatwo zauważyć jak uogólnić równania E-L na przypadek wielowymiarowy. Mamy dany funkcjonał w postaci:

$$I[q] = \int_a^b F(t, q_1(t), q_1'(t), \ldots, q_n(t), q_n'(t))x$$

określony na przestrzeni krzywych w ^{n}, różniczkowalnych, spełniających $q(a) = A$, $q(b) = B$. Krzywa $q : [a, b] \mapsto^n$ jest ekstremalą wtedy, i tylko wtedy gdy spełnia układ równań różniczkowych:

$$\forall\, i = 1, \ldots n \quad \frac{\partial F}{\partial q_i} - \frac{}{x}\frac{\partial F}{\partial q_i'} = 0$$

Literatura:

- "Rachunek wariacyjny" - I.M. Gelfand, S.W. Fomin