1DE2116:A - Matematyka - Metody optymalizacji EP
Zadanie izoperymetryczne.

Maciej Twardy

Materiał przygotowany w ramach projektu „NERW 2 PW. Nauka – Edukacja – Rozwój - Współpraca”. Projekt współfinansowany przez Unie Europejska w ramach Europejskiego Funduszu Społeczego. Program Operacyjny Wiedza Edukacja Rozwój 2014-2020, Os priorytetowa III Szkolnictwo Wyzsze dla gospodarki i rozwoju, Działanie 3.5 Kompleksowe programy szkół wyższych.

1 Wstęp

Rozpatrywane w ćwiczeniu zadanie, opracowane na podstawie [1, 2], jest ciekawe samo w sobie, jednak ma jeszcze tę dodatkową (a może nawet główną) zaletę, że ilustruje szeroką klasę metod rozwiązywania zadań „ciągłych” poprzez ich dyskretyzację.

Weźmy pod uwagę problem wyznaczenia krzywej na płaszczyźnie XY w taki sposób, aby powierzchnia zawarta między tą krzywą a osią X była możliwie największa. Z założenia będziemy rozpatrywać jedynie krzywe postaci

$$\mathcal{C} = \{(x, y) \,|\, y = f(x)\}\,, \tag{1}$$

gdzie $f : [0, a] \to \mathbb{R}$. Oznacza to, że dla danej wartości zmiennej x istnieje tylko jeden y, wyznaczony przez $y = f(x)$. Zakładamy również, że

$$f(0) = f(a) = 0. \tag{2}$$

Długość rozpatrywanej krzywej nie może przekraczać zadanej wartości L, czyli

$$\int_0^a \sqrt{1 + f'(x)^2}\mathrm{d}x \leqslant L. \tag{3}$$

Zagrodzona powierzchnia A jest równa

$$A = \int_0^a f(x)\mathrm{d}x. \tag{4}$$

Dodatkowo, żądamy aby rozpatrywana krzywa przechodziła przez pewne zadane punkty, tzn. dla pewnych ustalonych $\{t_1, \ldots, t_j\}$ oraz $\{\bar{y}_1, \ldots, \bar{y}_j\}$, mamy $f(x = t_k) = \bar{y}_k$, dla $k = 1, \ldots, j$. Ponadto nakładamy ograniczenia na maksymalną krzywiznę

$$|f''(x)| \leqslant C, \tag{5}$$

gdzie C jest pewną ustaloną stałą.
Posumowując, rozpatrywane zadanie jest postaci

$$\begin{aligned}
&\text{maximize} && \int_0^a f(x)\mathrm{d}x && \text{(6a)}\\
&\text{subject to} && \int_0^a \sqrt{1 + f'(x)^2}\mathrm{d}x \leqslant L && \text{(6b)}\\
& && |f''(x)| \leqslant C && \text{(6c)}\\
& && f(x = t_k) = \bar{y}_k, \qquad k = 1, \ldots, j && \text{(6d)}
\end{aligned}$$

dla pewnych ustalonych $\{t_1, \ldots, t_j\}$ oraz $\{\bar{y}_1, \ldots, \bar{y}_j\}$.

Ażeby sformułować tak postawione zadanie jako zadanie optymalizacji skończenie wymiarowej (ang. *finite dimensional optimization problem*), dyskretyzujemy zmienną x. Kładziemy $x_i = (i - 1)h$, $i = 1, \ldots, N + 1$, gdzie $h = a/N$ jest krokiem dyskretyzacji, mamy $x_1 = 0$, $x_{N+1} = Nh = a$. Oznaczamy $y_i = f(x_i)$, wówczas

$$\int_0^a f(x)\mathrm{d}x \approx \sum_{i=1}^{N} f(x_i)h = h\sum_{i=1}^{N} y_i. \tag{7}$$

Wyrażenie

$$h\sum_{i=1}^{N} y_i \tag{8}$$

jest zatem funkcją celu.
Ograniczenie długości krzywej można napisać w postaci

$$h\sum_{i=1}^{N}\sqrt{1 + \left(\frac{y_{i+1} - y_i}{h}\right)^2} \leqslant L, \tag{9}$$

gdzie przyjmujemy, że

$$f'(x_i) \approx \frac{y_{i+1} - y_i}{h}. \tag{10}$$

Oznaczając $g(x) = \sqrt{1 + f'(x)^2}$, możemy napisać

$$\begin{aligned}
\int_0^a \sqrt{1 + f'(x)^2}\mathrm{d}x &= \int_0^a g(x)\mathrm{d}x\\
&\approx h\sum_{i=1}^{N} g(x_i)\\
&= h\sum_{i=1}^{N}\sqrt{1 + f'(x_i)^2}\\
&\approx h\sum_{i=1}^{N}\sqrt{1 + \left(\frac{y_{i+1} - y_i}{h}\right)^2}\\
&= \sum_{i=1}^{N}\sqrt{h^2 + (y_{i+1} - y_i)^2}\\
&= \sum_{i=1}^{N}\left\|\begin{bmatrix} h\\ y_{i+1} - y_i\end{bmatrix}\right\|
\end{aligned} \tag{11}$$

Zauważmy, że wyrażenie $\sqrt{h^2 + (y_{i+1} - y_i)^2}$ jest długością odcinka łączącego punkty (x_i, y_i) i (x_{i+1}, y_{i+1}).

Dodatkowo, żądamy aby rozpatrywana krzywa przechodziła przez pewne zadane punkty. Oznaczmy zbiór indeksów tych punktów przez

$$\mathcal{F} \subseteq \{1, \ldots, N + 1\}\,. \tag{12}$$

Przyjmujemy, że dla $j \in \mathcal{F}$ mamy $y_j = y_j^{\text{fixed}}$, gdzie $\boldsymbol{y}^{\text{fixed}} \in \mathbb{R}^{N+1}$ (elementy wektora $\boldsymbol{y}^{\text{fixed}}$ których indeksy nie należą do $\mathcal{F}$ ignorujemy). Występującą w ograniczeniu krzywizny

$$|f''(x)| \leqslant C \tag{13}$$

drugą pochodną przybliżamy ilorazem

$$f''(x_i) \approx \frac{f'(x_{i+1}) - f'(x_i)}{h} \tag{14}$$

skąd otrzymujemy warunek

$$-C \leqslant \frac{y_{i+2} - 2y_{i+1} + y_i}{h^2} \leqslant C, \quad i = 1, \ldots, N-1. \tag{15}$$

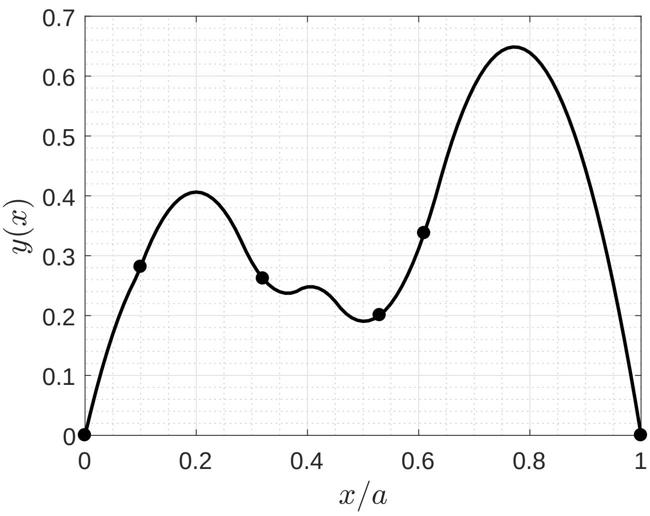
Podsumowując, należy rozwiązać następujące zadanie optymalizacji

$$\begin{aligned}
\text{maximize} \quad & h \sum_{i=1}^{N} y_i && \text{(16a)}\\
\text{subject to} \quad & \sum_{i=1}^{N} \left\| \begin{bmatrix} h \\ y_{i+1} - y_i \end{bmatrix} \right\| \leqslant L && \text{(16b)}\\
& \left| \frac{y_{i+2} - 2y_{i+1} + y_i}{h^2} \right| \leqslant C, \;\; i = 1, \ldots, N-1 && \text{(16c)}\\
& y_1 = 0 && \text{(16d)}\\
& y_{N+1} = 0 && \text{(16e)}\\
& y_j = y_j^{\text{fixed}}, \quad j \in \mathcal{F} && \text{(16f)}
\end{aligned}$$

względem zmiennych $y_1, \ldots, y_{N+1}$, które jest zadaniem wypukłym.

2 Zadania

Zadanie 1. Należy wyznaczyć poszukiwaną krzywą (tzn. punkty $y_1, \ldots, y_{N+1}$) dla danych z pliku `isoPerimData.mat` oraz wygenerować odpowiedni wykres (Rys. 1).



Rysunek 1: Wykres funkcji $f(x)$ maksymalizującej całkę $\int_0^a f(x)\mathrm{d}x$ przy zadanych ograniczeniach.

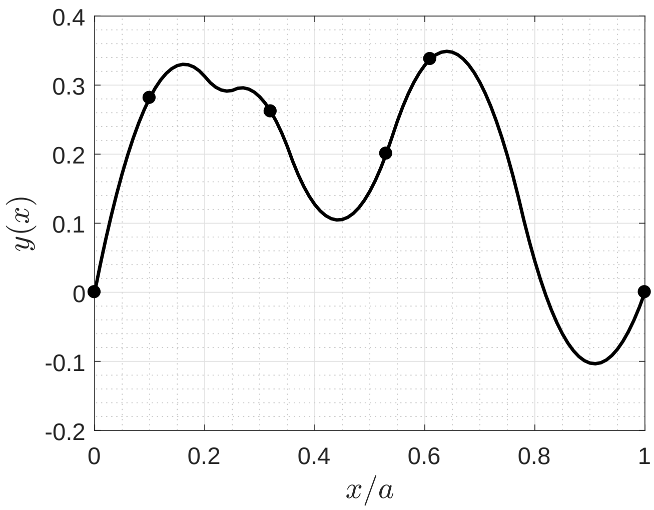
Modyfikacje zadania
Rozpatrzyć zadanie:

(a) *minimalizacji* pola pod krzywą (Rys. 2),
[Odp. $A = 0.1746$]

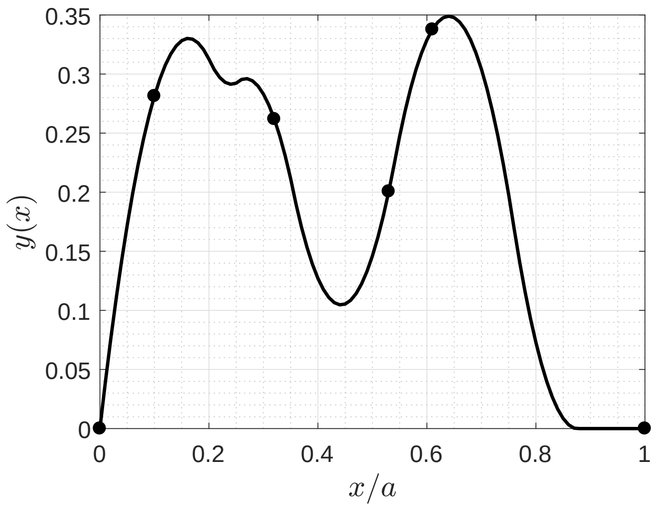
(b) minimalizacji pola pod krzywą przy nieujemnych zmiennych optymalizacyjnych (Rys. 3),
[Odp. $A = 0.1890$]

(c) maksymalizacji pola pod krzywą przy usunięciu ograniczenia na maksymalną krzywiznę.
[Odp. $A = 0.6016$]

Zinterpretuj otrzymane wyniki.



Rysunek 2: Wykres funkcji $f(x)$ minimalizującej całkę $\int_0^a f(x)\mathrm{d}x$ przy zadanych ograniczeniach.



Rysunek 3: Wykres funkcji $f(x)$ minimalizującej całkę $\int_0^a f(x)\mathrm{d}x$ przy zadanych ograniczeniach oraz dodatkowym warunku nieujemności zmniennych optymalizacyjnych.

Ponieważ zadania rozwiązujemy korzystając z pakietu CVX [3], należy zwrócić uwagę na następujące kwestie. Możemy zacytować następujący fragment [3]:

> As another example consider the function $\sqrt{x^2+1} = \left\| \begin{bmatrix} x \\ 1 \end{bmatrix} \right\|_2$, which is convex. If it is written as
>
> ```
> norm( [ x; 1 ] )
> ```
>
> (assuming `x` is a scalar variable or affine expression) it will be recognized by CVX as a convex expression, and therefore can be used in (appropriate) constraints and objectives. But if it is written as
>
> ```
> sqrt( x^2 + 1 )
> ```
>
> CVX will reject it, since convexity of this function does not follow from the CVX ruleset.

Drugą kwestią na którą trzeba zwrócić uwagę jest różnica między zmiennymi optymalizacyjnymi (ang. *optimization variables*) a zmiennymi wyrażeniowymi (ang. *expression holders*).

Szczegółową dyskusję na ten temat można znaleźć w [3] w podrozdziale 4.8 Assignment and expression holders. Tutaj podamy krótki przykład. Otóż załóżmy, że chcemy zminimalizować sumę kwadratów elementów pewnego wektora $x \in \mathbb{R}^n$, czyli

$$\sum_{i=1}^{n} x_i^2,$$

który jest zmienną optymalizacyjną, przy ograniczeniu $Ax \leqslant b$, gdzie macierz A i wektor b mają odpowiednie wymiary i są znane. Oczywiście możemy to zrobić za pomocą kodu

```
cvx_begin
   variable x(n)
   minimize ( x'*x )
   subject to
      A*x <= b
cvx_end
```

Możemy jednak skorzystać ze zmiennych wyrażeniowych i napisać następujący kod

```
cvx_begin
   variable x(n)
   expression s
   s = 0;
   for k = 1:n
      s = s + (x(k))^2;
   end
   minimize ( s )
   subject to
      A*x <= b
cvx_end
```

Oczywiście pierwszy kod jest krótszy i bardziej przejrzysty, zatem możemy powiedzieć, że jest lepszy. Jednak w wielu sytuacjach zmienne wyrażeniowe mogą być bardzo użyteczne, zadanie izoperymetryczne postaci (16) jest tutaj dobrym przykładem.

Literatura

[1] Stephen Boyd and Lieven Vandenberghe. *Convex Optimization.* Cambridge University Press, New York, NY, USA, 2004. `http://web.stanford.edu/~boyd/cvxbook/`.

[2] Stephen Boyd and Lieven Vandenberghe. *Additional Exercises for Convex Optimization.* 2004. `https://web.stanford.edu/~boyd/cvxbook/bv_cvxbook_extra_exercises.pdf`.

[3] Inc. CVX Research. CVX: Matlab software for disciplined convex programming, version 2.0. `http://cvxr.com/cvx`, August 2012.