Optymalizacja wypukła: wybrane zagadnienia

Michał Przyłuski

4 marca 2009

Ogólne zadanie optymalizacji

$$\begin{array}{rl} \text{minimalizuj} & f_0(x) \\ \text{p. o.} & f_i(x) \leqslant 0, \;\; i = 1, \dots, m, \\ & h_i(x) = 0, \;\; i = 1, \dots, p; \end{array}$$

- x — wektor zmiennych decyzyjnych,
- f_0 — funkcja celu,
- funkcje f_i oraz h_i — ograniczenia nierównościowe i równościowe.

Pewne klasy zadań optymalizacji

Rozwiązując takie zadania natrafiamy na szereg trudności, dlatego wolimy analizować ich szczególne postacie:

- zadania programowania liniowego — wszystkie funkcje są afiniczne,
- zadania programowania wypukłego — ograniczenia równościowe są afiniczne, pozostałe funkcje są wypukłe.

Wszystkie te problemy można (wydajnie) rozwiązywać metodą punktu wewnętrznego.

Stożkowe zadania programowania liniowego

$$\begin{aligned} \text{minimalizuj} \quad & f^T x \\ \text{p. o.} \quad & Ax + b \leqslant_K 0 \end{aligned}$$

- K — ustalony stożek dodatni w $\mathbb{R}^m$ (tj. generujący porządek $\leqslant_K$),
- wypukła $f \in \mathbb{R}^n$, $A \in \mathbb{R}^{m \times n}$ oraz $b \in \mathbb{R}^m$ — ustalone parametry zadania.

Szczególne przypadki to:

- programowanie stożkowe drugiego stopnia (SOCP),
- programowanie półokreślone (SDP).

Zadania SOCP

$$\begin{aligned} \text{minimalizuj} \quad & f^T x \\ \text{p. o.} \quad & \|A_i x + b_i\| \leqslant c_i^T x + d_i, \;\; i = 1, 2, \ldots, N; \end{aligned}$$

- $f \in \mathbb{R}^n$, $A_i \in \mathbb{R}^{(n_1-1)\times n}$, $b_i \in \mathbb{R}^{n_i-1}$, $c_i \in \mathbb{R}^n$ oraz $d_i \in \mathbb{R}$ — ustalone parametry zadania.

Ograniczenia te określają stożek drugiego stopnia, bowiem nierówność

$$\|A_i x + b_i\| \leqslant c_i^T x + d_i$$

jest spełniona wtedy i tylko wtedy, gdy

$$\begin{bmatrix} A_i \\ c_i^T \end{bmatrix} x + \begin{bmatrix} b_i \\ d_i \end{bmatrix} \in C_{n_i},$$

gdzie C_{n_i} stożek drugiego stopnia wymiaru n_i.

Zadania SDP

Niech

$$F(x) := F_0 + \sum_{i=1}^{m} x_i F_i,$$

gdzie $F_0, F_1, \ldots, F_m$ — ustalone macierze symetryczne.

$$\begin{array}{rl} \text{minimalizuj} & c^T x \\ \text{p. o.} & F(x) \succcurlyeq 0 \end{array}$$

- wektor $x = (x_1, x_2, \ldots, x_n)^T$ — zmienna decyzyjna,
- macierze F_i oraz wektor c — parametry zadania,
- warunek dodatniej półokreśloności macierzy $F(x)$ — ograniczenie (LMI).

Do postaci SOCP lub SDP możemy sprowadzić m. in.:

- zadania programowania geometrycznego,
- QCQP,
- (krzepkie) zadania programowania liniowego,
- krzepkie zadania najmniejszych kwadratów,
- zadania minimalizacji normy.

Minimalizacja normy (1)

Celem jest aproksymacja (w ℓ_1) zadanego wektora $b \in \mathbb{C}^k$ wektorami z podprzestrzeni liniowej $\mathbb{C}^k$ będącej obrazem zadanej macierzy $A \in \mathbb{C}^{k \times q}$.
Szukamy zmiennej decyzyjnej $x \in \mathbb{C}^q$, która jest rozwiązaniem zadania:

$$\text{minimalizuj} \quad \|Ax - b\|_1.$$

Niech z^{real} i z^{im} — to odpowiednio część rzeczywistą i urojoną z, a a_i — i-ty wiersz macierzy A.
Minimalizacja powyższej normy to minimalizacja sumy k składników postaci $|a_i x - b_i|$, tj.

$$\left((a_i^{\text{real}} x^{\text{real}} - a_i^{\text{im}} x^{\text{im}} - b_i^{\text{real}})^2 + (a_i^{\text{real}} x^{\text{im}} + a_i^{\text{im}} x^{\text{real}} - b_i^{\text{im}})^2 \right)^{½}.$$

Minimalizacja normy (2)

Jest to norma (euklidesowa):

$$\begin{bmatrix} a_i^{\text{real}} & -a_i^{\text{im}} \\ a_i^{\text{im}} & a_i^{\text{real}} \end{bmatrix} \begin{bmatrix} x^{\text{real}} \\ x^{\text{im}} \end{bmatrix} - \begin{bmatrix} b^{\text{real}} \\ b^{\text{im}} \end{bmatrix}.$$

Sformułowane zadanie można przedstawić w postaci SOCP:

$$\begin{array}{ll} \text{minimalizuj} & \sum_{i=1}^{k} t_i \\ \text{p. o.} & \left\| \begin{bmatrix} a_i^{\text{real}} & -a_i^{\text{im}} \\ a_i^{\text{im}} & a_i^{\text{real}} \end{bmatrix} \begin{bmatrix} x^{\text{real}} \\ x^{\text{im}} \end{bmatrix} - \begin{bmatrix} b^{\text{real}} \\ b^{\text{im}} \end{bmatrix} \right\| \leqslant t_i, \quad i = 1, 2, \ldots, k, \end{array}$$

- liczby t_i oraz wektory x^{real} i x^{im} — zmienne decyzyjne.

Sformułowanie problemu — układy anten

- Ustalona liczba anten składowych n,
- ustalone rozmieszczenie na płaszczyźnie w (x_i, y_i),
- poszukiwany jest zestaw wag $(w_1, \dots, w_n) \in \mathbb{C}$.

Cel optymalizacji: maksymalna czułość w zadanym kierunku (θ_t) i minimalna czułość poza wiązką.
Dla kąta padania θ sygnał informacyjny z i-tej anteny: $e^{j(x_i \cos\theta + y_i \sin\theta)}$.
Wkład i-tej anteny do całkowitego odebranego sygnału:
$y_i(\theta) = w_i e^{j(x_i \cos\theta + y_i \sin\theta)} =$
$(w_i^{\mathrm{real}} \cos\gamma_i(\theta) - w_i^{\mathrm{im}} \sin\gamma_i(\theta)) + j(w_i^{\mathrm{real}} \sin\gamma_i(\theta) + w_i^{\mathrm{im}} \cos\gamma_i(\theta))$,
gdzie $\gamma_i(\theta) = x_i \cos\theta + y_i \sin\theta$.

Problem ciągły

$$
\begin{array}{ll}
\text{minimalizuj} & \max\limits_{|\theta-\theta_t|>\Delta} |y(\theta)| \\
\text{p. o.} & y(\theta_t) = 1
\end{array}
$$

- Sygnał informacyjny odebrany przez układ anten to $y(\theta) = \sum_{i=1}^{n} y_i(\theta)$,
- $y(\theta)$ jest liniową funkcją wektora wag w,
- zmienne decyzyjne to wagi w_i, zawarte w funkcji $y(\theta)$.

Poprzez dyskretyzację kąta θ zagadnienie można przybliżyć, aby stało się zadaniem SOCP.

Problem zdyskretyzowany

Oryginalny problem można aproksymować zadaniem:

$$\begin{array}{rll} \text{minimalizuj} & t & \\ \text{p. o.} & |y(\theta_i)| \leqslant t, & \text{dla } |\theta_i - \theta_k| > \Delta, \\ & y(\theta_k) = 1, & \end{array}$$

będącym w postaci SOCP, gdy wyrazimy je w terminach części rzeczywistych i urojonych zmiennych oraz danych.

Sformułowanie problemu — siła chwytająca

Bryła sztywna trzymana przez N palców robota, w punktach $r^1, r^2, \ldots, r^N \in \mathbb{R}^3$.
Siła $\vec{F}^i$ może być rozłożona na dwie składowe:

- $(\vec{n}^i)^T \vec{F}^i \vec{n}^i$ — normalną do powierzchni,
- $\vec{F}^i - (\vec{n}^i)^T \vec{F}^i \vec{n}^i$ — styczną do powierzchni.

Składowa styczna — tarcie statyczne między robotem a bryłą — wartość ograniczona przez iloczyn μ ze składową normalną

$$\|\vec{F}^i - (\vec{n}^i)^T \vec{F}^i \vec{n}^i\| = \|(I - \vec{n}^i (\vec{n}^i)^T) \vec{F}^i\| \leqslant \mu (\vec{n}^i)^T \vec{F}^i, \;\; i = 1, 2, \ldots, N.$$

Warto zauważyć, że względem współrzędnych $\vec{F}^i$ są to ograniczenia drugiego stopnia.

Stabilny uchwyt

Na bryłę sztywną działają pewne zewnętrzne siły oraz momenty sił; ich wypadkowe to $\vec{F}^{\text{ext}}$ oraz $\vec{M}^{\text{ext}}$.
Szukamy sił $\vec{F}^i$, aby

- $\sum_{i=1}^{N} \vec{F}^i + \vec{F}^{\text{ext}} = 0$,
  $\sum_{i=1}^{N} \vec{r}^i \times \vec{F}^i + \vec{M}^{\text{ext}} = 0$
  — równowaga statyczna sił i momentów działających na bryłę,
- $\|(I - \vec{n}^i(\vec{n}^i)^T)\vec{F}^i\| \leqslant \mu(\vec{n}^i)^T\vec{F}^i$ — tarcie statyczne,
- $(\vec{n}^i)^T\vec{F}^i \leqslant f_{\max}$ — limit na składowe normalne.

Najdelikatniejszy uchwyt

Problem wyboru sił $\vec{F}^i$:

- gwarantujących stabilny uchwyt,
- minimalizować największą składową normalną do powierzchni.

$$
\begin{array}{ll}
\text{minimalizuj} & t \\
\text{p. o.} & (\vec{n}^i)^T \vec{F}_i \leqslant t, \\
& \sum_{i=1}^{N} \vec{F}^i + \vec{F}^{\text{ext}} = 0, \quad \sum_{i=1}^{N} \vec{r}^i \times \vec{F}^i + \vec{M}^{\text{ext}} = 0, \\
& \|(I - \vec{n}^i (\vec{n}^i)^T)\vec{F}^i\| \leqslant \mu (\vec{n}^i)^T \vec{F}^i, \quad \text{dla } i = 1, 2, \ldots, N,
\end{array}
$$

Gdy ograniczenia związane z tarciem zostaną przybliżone nierównościami liniowymi, problem ten stanie się ZPL.

Wnioski

Podstawowe przesłanki za stosowaniem optymalizacji wypukłej:

- wiele zadań można przekształcić do takiej postaci
  — zastosowania w licznych dziedzinach,
- wydajne algorytmy rozwiązywania zadań.