Notes

Kopnov Alexandr

22 мая 2023 г.

Содержание

1 Постановка

Решить задачу

$$\min \phi_0(x) \quad \forall x \in \Omega$$

$$\Omega = \{x \in \mathbb{R}^n | \phi_i(x) \leqslant 0, i = \overline{1, m}\}$$

методом отсекающей гиперплоскости. Обосновать сходимость метода.

1.1 Данные задачи:

В нашем случае $\phi_0 = x_1^2 + 4x_2^2 + \sin(6x_1 + 7x_2) + 3x_1 + 2x_2$ — нелинейная функция, значит вводим $\phi_4(x_1, x_2, x_3) = \phi_0 - x_3$, а в качестве функции цели выберем $\phi_0(x) = x_3$ Из решения задачи без ограничений знаем точку минимума:

$$([-1.9043886748215388, -0.36794672198200706], -3.2716974195454807)$$

По условиям применимости метода $\phi_i(x), i = \overline{1, m}$ должны быть выпуклыми. Рассмотрим выпуклость функции ϕ_4

$$\phi_4(x) = \phi_{41} + \phi_{42} - x_3$$

$$\phi_{41} = x_1^2 + 4x_2^2 + 3x_1 + 2x_2 \text{ — выпуклая}$$

$$-x_3 \text{ — линейная} \implies \text{выпуклая}$$

$\phi_{42} = \sin(6x_1 + 7x_2)$

Для поиска интервала на котором функция ϕ_{42} будет выпуклой, воспользуемся теоремой: Пусть S — непустое открытое выпуклое мн-во, $\phi(x)$ — дважды дифференцируемая функция. Тогда для того, чтобы $\phi(x)$ была выпуклой функцией на S необходимо и достаточно, чтобы её гессиан $H(x)$ был положительно-полуопределённой матрицей

$$H_{42}(x) = \begin{pmatrix} -36\sin(6x_1 + 7x_2) & -42\sin(6x_1 + 7x_2) \\ -42\sin(6x_1 + 7x_2) & -49\sin(6x_1 + 7x_2) \end{pmatrix}$$

Матрица будет положительно полуопределённой при $\sin(6x_1+7x_2) \leqslant 0$. (изначально исследование проводилось для всей функции, а не только для синуса, но это не внесло значительных изменений $(\sin(6x_1 + 7x_2) \leqslant 0.0414)$, хотя привело к существенному удлинению выкладок)

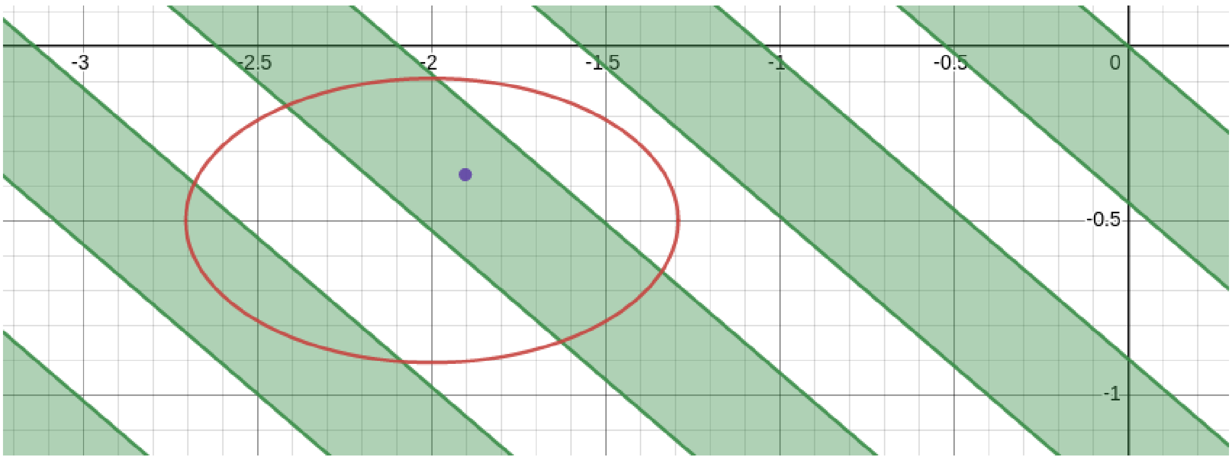
Т.о. поставим линейные условия:

$$\begin{aligned} 6x_1 + 7x_2 &\geqslant -5\pi \\ 6x_1 + 7x_2 &\leqslant -4\pi \end{aligned} \implies \begin{aligned} -6x_1 - 7x_2 - 5\pi &\leqslant 0 \\ 6x_1 + 7x_2 + 4\pi &\leqslant 0 \end{aligned}$$

Поставим нелинейное условие:

$$(x+2)^2 + 3(y+0.5)^2 \leqslant 0.5 \implies (x+2)^2 + 3(y+0.5)^2 - \frac{1}{2} \leqslant 0$$

Так выглядит область допустимых значений при отображении через `Desmos`



В качестве начального приближения возьмём многогранное мн-во $S_0 = \{x | a_i^T x - b_i \leqslant 0, i = \overline{-3, 0}\}$ где

$$\begin{aligned} 6x_1 + 7x_2 &\leqslant -4\pi \\ -6x_1 - 7x_2 &\leqslant 5\pi \\ x_2 &\leqslant -0.05 \\ x_2 + 0.5 &\geqslant x_1 \\ -x_1 - x_2 + x_3 &\geqslant -2 \\ x_3 &\leqslant 50 \end{aligned} \implies \begin{aligned} 6x_1 + 7x_2 &\leqslant -4\pi \\ -6x_1 - 7x_2 &\leqslant 5\pi \\ x_2 &\leqslant -0.05 \\ x_1 - x_2 &\leqslant 0.5 \\ x_1 + x_2 - x_3 &\leqslant 2 \\ x_3 &\leqslant 50 \end{aligned}$$

Представление исходного полиэдра в виде матрицы ограничений

```
6,3
0,0,1
6,7,0,LT,-12.566
-6,-7,0,LT,15.708
0,1,0,LT,-0.05
1,-1,0,LT,0.5
1,1,-1,LT,2
0,0,1,LT,50
1,2,3
```

1.2 О субградиенте

Вектор $y \in \mathbb{R}^n$ называется субградиентом функции $\phi(x)$ в точке $x_0 \in S$, если

$$\phi(x) \geqslant \phi(x_0) + y^T(x - x_0), \quad \forall x \in S$$

Заметим, что $y = y(x_0)$
В прошлом семестре была доказана теорема:
Пусть $\phi(x)$ — выпукло-дифференцируемая функция, заданная на выпуклом множестве S, причём $\dot{S} \neq \emptyset$. Тогда в каждой точке $x_0 \in \dot{S}$ субградиент $y = \nabla\phi(x_0)$

Рассмотрим функцию $\phi(x) = \max_i\{\phi_i(x)\}$. Запишем субградиент функции $\phi_i(x)$

$$\phi_i(x) \geqslant \phi_i(x_0) + y_i^T(x - x_0)$$

По теореме, $y_i = \nabla\phi_i(x_0)$. Будем использовать $\phi_i : \phi(x_0) = \phi_i(x_0)$ (наибольшую из функций в заданной точке). Тогда

$$\phi(x) \geqslant \phi_i(x) \geqslant \phi_i(x_0) + \nabla\phi_i^T(x - x_0) = \phi(x_0) + \nabla\phi_i^T(x - x_0)$$

Таким образом, $\nabla\phi_i(x_0)$ — субградиент функции ϕ в т. x_0

2 Условие оптимальности

В общем случае необходимым условием условного экстремума является равенство нулю градиента функции Лагранжа:

$$L(x, \lambda) = \phi_0(x) + \sum_{i=1}^{n} \lambda_i \phi_i(x)$$

Где $\phi_i(x), i = \overline{1, n}$ — функции, задающие уравнения связи $\phi_i = 0$. Функция L является функцией не только x, но и λ, поэтому для того, чтобы проверить условие на градиент, нужно получить λ. Однако исследуемый метод не находит λ, что означает, что мы не можем использовать это условие.

Функция Лагранжа для нашей задачи:

$$L(x, y) = x_1^2 + 4x_2^2 + \sin(6x_1 + 7x_2) + 3x_1 + 2x_2 + \qquad y_1\left((x_1 + 2)^2 + 3(y + 0.5)^2 - \frac{1}{2}\right) +$$
$$+ y_2(-6x_1 - 7x_2 - 5\pi) + y_3(6x_1 + 7x_2 + 4\pi)$$

Не знаю надо ли писать список с учётом последнего параграфа.

1. $\nabla L(x, y) = 0$

$$\begin{cases} 2x_1 + 6\cos(6x_1 + 7x_2) + 3 + 2y_1(x_1 + 2) - 6y_2 + 6y_3 \\ 8x_2 + 7\cos(6x_1 + 7x_2) + 2 + 6y_1(x_2 + 0.5) - 7y_2 + 7y_3 \end{cases}$$

2. Условия дополняющей нежесткости:

$$\begin{cases} y_1\left((x_1 + 2)^2 + 3(y + 0.5)^2 - \frac{1}{2}\right) = 0 \\ y_2(-6x_1 - 7x_2 - 5\pi) = 0 \\ y_3(6x_1 + 7x_2 + 4\pi) \end{cases}$$

3. Ограничение на знак:

$$\begin{cases} y_1 \geqslant 0 \\ y_2 \geqslant 0 \\ y_3 \geqslant 0 \end{cases}$$

Точка экстремума находится внутри множества, поэтому все ограничения выполняются пассивно и $y = 0$. Таким образом, градиент функции Лагранжа равен градиенту функции ϕ_0.

При $\epsilon = 10^{-3}$, $\|x - x_*\| = 10^{-3}, \|\nabla\phi_0(x)\| = 4 \cdot 10^{-3}$