Методы оптимизации, HW4

Иванов Вячеслав, группа 699

7 ноября 2018 г.

Задача 1. Решить, используя условия ККТ:

$$\begin{aligned} &\min_{(x,y)} (x-3)^2 - (y-2)^2 \\ \text{s.t. } & y = x+1 \\ & y \leqslant -x+3 \end{aligned}$$

Решение. Перепишем целевую функцию, используя условие $y = x+1$:

$$(x-3)^2 - (y-2)^2 \to (x-3)^2 - (x-1)^2 = 4(2-x)$$

Т.е. на множестве допустимых решений функция аффинная, а потому и выпуклая, и вогнутая. Следовательно, условия ККТ являются достаточными:

$$\begin{aligned} &\mathcal{L}(x, y, \lambda, \mu) = ((x-3)^2 - (y-2)^2) + \lambda(-x+y-1) + \mu(x+y-3) \\ &y_* = x_* + 1 \\ &y_* \leqslant -x_* + 3 \\ &\mu_* \geqslant 0 \\ &\mu_*(y_* + x_* - 3) = 0 \\ &\nabla_{(x,y)}\mathcal{L}(x_*, y_*, \lambda_*, \mu_*) = \begin{bmatrix} 2(x_*-3) - \lambda_* + \mu_* \\ -2(y_*-2) + \lambda_* + \mu_* \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \end{aligned}$$

Рассмотрим два случая, возникающих из условий вспомогательной жёсткости:

1. $\mu_* = 0$:

$$\begin{cases} 2(x_*-3) - \lambda_* &= 0, \\ 2(y_*-2) + \lambda_* &= 0, \\ y_* = x_* + 1 \end{cases}$$

$$\begin{cases} \lambda_* &= 2(x_*-3), \\ \lambda_* &= 2(x_*-1), \\ y_* &= x_* + 1 \end{cases}$$

Из первых двух уравнений системы видно, что решений нет.

2. $\mu_* > 0$:

$$\begin{cases} 2(x_* - 3) - \lambda_* + \mu_* & = 0, \\ 2(y_* - 2) + \lambda_* + \mu_* & = 0, \\ y_* = x_* + 1 \\ y_* = -x_* + 3 \end{cases}$$
$$\begin{cases} \lambda_* = 2(x_* - 3) + \mu_*, \\ \mu_* = -(y_* - 2) - (x_* - 3), \\ x_* = y_* - 1, \\ y_* = 2 \end{cases}$$
$$\begin{cases} \lambda_* = -2, \\ \mu_* = 2, \\ x_* = 1, \\ y_* = 2 \end{cases}$$

Значит, $(1, 2)$ — решение задачи оптимизации.

□

Задача 2. Для задачи

$$\min_{x \in \mathbb{R}^3} \; x_1^2 + 2x_2^2 + x_3$$
$$\text{s.t. } |x_1 - 2x_2 + 3x_3| \leqslant 4$$

найти множество стационарных точек и проверить, являются будут ли они решениями или являются седловыми.

Решение.
Матрица Гессе целевой функции имеет вид $\mathrm{diag}(2, 4, 0)$, в связи с чем она является глобально выпуклой по второму дифференциальному критерию. Т.к. условия аффинные, задача выпуклая и условия ККТ являются достаточными.
Стационарная точка задачи условной оптимизации — точка $(\mathbf{x}_*, \lambda_*, \mu_*)$, т.ч.

$$\mathcal{L}(\mathbf{x}_*, \lambda_*, \mu_*) = \mathbf{0}$$

В канонической постановке задача примет вид:

$$\begin{aligned} \min_{x \in \mathbb{R}^3} \; & x_1^2 + 2x_2^2 + x_3 \\ \text{s.t. } & -4 + (x_1 - 2x_2 + 3x_3) \leqslant 0 \\ & -4 - (x_1 - 2x_2 + 3x_3) \leqslant 0 \end{aligned}$$

Отсюда:

$$\mathcal{L}(\mathbf{x}, \mu) = x_1^2 + 2x_2^2 + x_3 + \mu_1(-4 + (x_1 - 2x_2 + 3x_3)) + \mu_2(-4 - (x_1 - 2x_2 + 3x_3))$$

Тогда рассмотрим все четыре случая, допускаемые условиями ККТ:

1. $\mu_1 > 0,\ \mu_2 > 0$:

$$\nabla\mathcal{L}(\mathbf{x}, \mu) = \begin{bmatrix} 2x_1 + \mu_1 - \mu_2 \\ 4x_2 - 2\mu_1 + 2\mu_2 \\ 1 + 3\mu_1 - 3\mu_2 \\ -4 + x_1 - 2x_2 + 3x_3 \\ -4 - x_1 + 2x_2 - 3x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\begin{cases} x_1 = \frac{1}{2}\left(\mu_2 - \mu_1\right), \\ x_2 = -\frac{1}{2}\left(\mu_2 - \mu_1\right), \\ \mu_2 - \mu_1 = \frac{1}{3}, \\ x_3 = \frac{1}{3}(4 - x_1 + 2x_2), \\ x_3 = \frac{1}{3}(-4 - x_1 + 2x_2) \end{cases}$$

$$\begin{cases} x_1 & = \frac{1}{6}, \\ x_2 & = -\frac{1}{6}, \\ \mu_2 - \mu_1 & = -\frac{1}{3}, \\ x_3 & = \frac{7}{6}, \\ x_3 & = -\frac{3}{2} \end{cases}$$

Из трёх последних уравнений системы видно, что при $\mu_1 > 0$ и $\mu_2 > 0$ **стационарных точек нет** (что, впрочем, геометрически очевидно). Тем не менее, матрица системы была осознанно приведена к наиболее простому виду, т.к. это пригодится в дальнейшем.

2. $\mu_1 > 0,\ \mu_2 = 0$:
Из условия $\mu_2 - \mu_1 = \frac{1}{3}$ видно, что такой случай не подходит.

3. $\mu_1 = 0,\ \mu_2 > 0$:

$$\begin{cases} x_1 & = \frac{1}{6}, \\ x_2 & = -\frac{1}{6}, \\ x_3 & = -\frac{3}{2}, \\ \mu_1 & = \frac{1}{3} \end{cases}$$

Т.е. $\left(\frac{1}{6}, -\frac{1}{6}, -\frac{3}{2}\right)$ — стационарная точка. Подстановкой её в первое ограничение убеждаемся, что она является решением.

4. $\mu_1 = 0,\ \mu_2 = 0$:
Из условия $\mu_2 - \mu_1 = \frac{1}{3}$ видно, что такой случай тоже не подходит.

□

Задача 3. Для задачи

$$\min_{\mathbf{x}\in\mathbb{R}^2} (x_1 - 2)^2 + (x_2 + 1)^2$$
$$\text{s.t. } \max(x_1 + x_2 + 1, x_1 - x_2 - 1) + 1 \leqslant 0$$

найти множество стационарных точек и классифицировать их на решения и седловые.

Решение. В каноническом виде задача запишется так:

$$\min_{\mathbf{x}\in\mathbb{R}^2} (x_1 - 2)^2 + (x_2 + 1)^2$$
$$\text{s.t. } x_1 + x_2 + 2 \leqslant 0$$
$$x_1 - x_2 \leqslant 0$$

Матрица Гессе целевой функции имеет вид $\mathrm{diag}(2,2)$, в связи с чем она является глобально выпуклой по второму дифференциальному критерию.
В силу аффинности ограничений, задача оптимизации сама является выпуклой.
Имеем:-

$$\mathcal{L}(\mathbf{x},\mu) = (x_1-2)^2 + (x_2+1)^2 + \mu_1(x_1+x_2+2) + \mu_2(x_1-x_2)$$

$$\nabla\mathcal{L}(\mathbf{x},\mu) = \begin{bmatrix} 2(x_1-2)+\mu_1+\mu_2 \\ 2(x_2+1)+\mu_1-\mu_2 \\ x_1+x_2+2 \\ x_1-x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\begin{cases} 2(x_1-2)+\mu_1+\mu_2 = 0, & ((1)-(2))/2 \\ 2(x_2+1)+\mu_1-\mu_2 = 0, & ((1)+(2))/2 \\ x_1+x_2+2 = 0, \\ x_1-x_2 = 0 \end{cases}$$

$$\begin{cases} \mu_2 = 3-x_1+x_2, \\ \mu_1 = 1-x_1-x_2, \\ x_1 = -2-x_2, \\ x_1 = x_2 \end{cases}$$

Рассмотрим теперь все четыре случая, допускаемые условиями ККТ:

1. $\mu_1 > 0,\ \mu_2 > 0$:

$$\begin{cases} \mu_1 = 1-x_1-x_2, \\ \mu_2 = 3-x_1+x_2, \\ x_1 = -2-x_2, \\ x_1 = x_2 \end{cases}$$

$$\begin{cases} \mu_1 = 3, \\ \mu_2 = 3, \\ x_1 = -1, \\ x_2 = x_1 \end{cases}$$

   Т.е. $\mathbf{x}^* = (-1,-1)$ — стационарная точка. В силу выпуклости задачи она является решением.

2. $\mu_1 > 0,\ \mu_2 = 0$:

$$\begin{cases} \mu_1 = 1-x_1-x_2, \\ 0 = 3-x_1+x_2, \\ x_2 = -2-x_1 \end{cases}$$

$$\begin{cases} \mu_1 = 1-x_1-x_2, \\ x_1 = \frac{1}{2}, \\ x_2 = -2-x_1 \end{cases}$$

$$\begin{cases} \mu_1 = 3, \\ x_1 = \frac{1}{2}, \\ x_2 = -\frac{5}{2} \end{cases}$$

Т.е. $x^* = \left(\frac{1}{2},\ -\frac{5}{2}\right)$ — стационарная точка. Подставим её во второе ограничение, чтобы проверить, выполнены ли условия ККТ:

$$x_1 - x_2 = \frac{1}{2} + \frac{5}{2} = 3 > 0$$

Следовательно, эта точка не является решением.

3. $\mu_1 = 0,\ \mu_2 > 0$:

$$\begin{cases} 0 = 1 - x_1 - x_2, \\ \mu_2 = 3 - x_1 + x_2, \\ x_1 = x_2 \end{cases}$$

$$\begin{cases} \mu_2 = 3 - x_1 + x_2, \\ x_1 = \frac{1}{2}, \\ x_2 = \frac{1}{2} \end{cases}$$

$$\begin{cases} \mu_2 = 3, \\ x_1 = \frac{1}{2}, \\ x_2 = \frac{1}{2} \end{cases}$$

Т.е. $x^* = \left(\frac{1}{2},\ \frac{1}{2}\right)$ — стационарная точка. Подставим её в первое ограничение, чтобы проверить, выполнены ли условия ККТ:

$$x_1 + x_2 + 2 = \frac{1}{2} + \frac{1}{2} + 2 > 0$$

Следовательно, эта точка не является решением.

4. $\mu_1 = 0,\ \mu_2 = 0$:

$$\begin{cases} 1 - x_1 - x_2 = 0, \\ 3 - x_1 + x_2 = 0 \end{cases}$$

$$\begin{cases} x_1 = 1 - x_2 \\ 3 - (1 - x_2) + x_2 = 0 \end{cases}$$

$$\begin{cases} x_1 = 2, \\ x_2 = -1 \end{cases}$$

Подстановка во второе ограничение показывает, что стационарная точка $\mathbf{x}^* = (2, -1)$ не является решением: $\mu_1 - \mu_2 = 3 > 0$.

□

Задача 4. Проверить задачу на выпуклость и решить её, используя условия ККТ.

$$\min_{\mathbf{x} \in S^2} -3x_1^2 + x_2^2 + 2x_3^2 + 2(x_1 + x_2 + x_3)$$

Решение. Функционал минимизируется на единичной сфере в $\mathbb{R}^3$, которая выпуклым множеством не является. Следовательно, поставлена задача невыпуклой оптимизации.
Запишем Лагранжиан:

$$\mathcal{L}(\mathbf{x}, \lambda) = -3x_1^2 + x_2^2 + 2x_3^2 + 2(x_1 + x_2 + x_3) + \lambda(x_1^2 + x_2^2 + x_3^2 - 1)$$

И приравняем его градиент к нулю, чтобы найти стационарные точки:

$$\nabla\mathcal{L}(\mathbf{x},\lambda) = \begin{bmatrix} -6x_1 + 2 + 2\lambda x_1 \\ 2x_2 + 2 + 2\lambda x_2 \\ 4x_3 + 2 + 2\lambda x_3 \\ x_1^2 + x_2^2 + x_3^2 - 1 \end{bmatrix} = \begin{bmatrix} 2(\lambda-3)x_1 + 2 \\ 2(\lambda+1)x_2 + 2 \\ 2(\lambda+2)x_3 + 2 \\ x_1^2 + x_2^2 + x_3^2 - 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\begin{cases} (\lambda-3)x_1 + 1 = 0, \\ (\lambda+1)x_2 + 1 = 0, \\ (\lambda+2)x_3 + 1 = 0, \\ x_1^2 + x_2^2 + x_3^2 - 1 = 0 \end{cases}$$

$$\begin{cases} x_1 = -\frac{1}{\lambda-3}, \\ x_2 = -\frac{1}{\lambda+1}, \\ x_3 = -\frac{1}{\lambda+2}, \\ x_1^2 + x_2^2 + x_3^2 - 1 = 0 \end{cases}$$

При аналитическом решении полученной системы уравнений пришлось бы решать уравнение шестой степени по λ. Чтобы этого избежать, было разрешено воспользоваться матпакетами. Я выбрал Wolfram Mathematica:

```
In[22]:= stationaryPoints = NSolve[{(λ - 3) x1 + 1 == 0, (λ + 1) x2 + 1 == 0, (λ + 2) x3 + 1 == 0, x1^2 + x2^2 + x3^2 - 1 == 0}, {x1, x2, x3, λ}, Reals]
         -3 x1^2 + x2^2 + 2 x3^2 + 2 (x1 + x2 + x3) /. stationaryPoints
Out[22]= {{x1 → -0.965973, x2 → -0.198601, x3 → -0.165694, λ → 4.03523}, {x1 → 0.16262, x2 → 0.46527, x3 → 0.870103, λ → -3.14929},
          {x1 → 0.902445, x2 → -0.345794, x3 → -0.256944, λ → 1.8919}, {x1 → 0.360167, x2 → -0.817321, x3 → -0.44974, λ → 0.223509}}
Out[23]= {-5.36549, 4.64728, -1.59219, -1.1304}
```

Рис. 1: Стационарные точки (Out[22]) и значения функционала на них (Out[23])

```
NMinimize[{-3 x1^2 + x2^2 + 2 x3^2 + 2 (x1 + x2 + x3), x1^2 + x2^2 + x3^2 == 1}, {x1, x2, x3}]
{-5.36549, {x1 → -0.965973, x2 → -0.198601, x3 → -0.165694}}
```

Рис. 2: Решение, найденное численными методами

Классифицируем стационарные точки с помощью достаточного условия условного экстремума.

Достаточные условия. Пусть функции $f(x)$, $\varphi_i(x)$, $i = 1, 2, \ldots$ $\ldots, m$, $x \in R^n$, дважды непрерывно дифференцируемы в окрестности точки x^0, и пусть в этой точке выполняются необходимые условия существования условного экстремума функции $f(x)$ при ограничениях (4).

Тогда, если при выполнении условий

$$d\varphi_i(x^0) = \sum_{k=1}^{n} \frac{\partial \varphi_i(x_0)}{\partial x_k}\, dx_k = 0, \quad \sum_{k=1}^{n} dx_k^2 > 0, \tag{9}$$

второй дифференциал $d^2L(x^0)$ функции Лагранжа является положительно (отрицательно) определенной квадратичной формой, то функция $f(x)$ в точке x^0 имеет условный строгий минимум (максимум).

§ 5. *Экстремумы функций* 115

Если при условиях (9) второй дифференциал $d^2L(x^0)$ является неопределенной квадратичной формой, то в точке x^0 условного экстремума нет.

Рис. 3: Кудрявцев, "Сборник задач по математическому анализу, Том 2. стр. 114-115

$$\nabla^2_{\mathbf{x}} L(\mathbf{x}, \lambda) = \begin{bmatrix} 2\lambda - 6 & 0 & 0 \\ 0 & 2\lambda + 2 & 0 \\ 0 & 0 & 2\lambda + 4 \end{bmatrix}$$

$$d\varphi(x) = x_1 dx_1 + x_2 dx_2 + x_3 dx_3 = 0$$

$$dx_3 = \frac{-x_1 dx_1 - x_2 dx_2}{x_3}$$

$$d^2L(\mathbf{x}, \lambda) = 2(\lambda - 3)dx_1^2 + 2(\lambda + 1)dx_2^2 + 2(\lambda + 2)dx_3^2$$

Поскольку стационарные точки получены численно и подставлять их в уравнения вручную было бы безумием, я написал небольшой скрипт на Wolfram Language, который делает это за меня:

```
In[83]:= f[x1_, x2_, x3_] := -3 x1^2 + x2^2 + 2 x3^2 + 2 (x1 + x2 + x3);
stationaryPoints = NSolve[{(λ - 3) x1 + 1 == 0, (λ + 1) x2 + 1 == 0, (λ + 2) x3 + 1 == 0, x1^2 + x2^2 + x3^2 - 1 == 0}, {x1, x2, x3, λ}, Reals];
For[i = 1, i ≤ 4, ++i,
 {
   (dx3 = (-x1 dx1 - x2 dx2)/x3 /. stationaryPoints[[i]];
   d2L = FullSimplify[2 (λ - 3) (dx1)^2 + 2 (λ + 1) (dx2)^2 + 2 (λ + 2) (dx3)^2 /. stationaryPoints[[i]]];
   L11 = d2L /. {dx2 -> 0, dx1 → 1};
   L12 = 1/2 (d2L /. {(dx1)^2 → 0, (dx2)^2 → 0, dx1 → 1, dx2 → 1});
   L21 = L12;
   L22 = d2L /. {dx1 → 0, dx2 → 1};
   d2LMx = {{L11, L12}, {L21, L22}};
   Print["Stationary point:", stationaryPoints[[i]]];
   Print["Function value at this point: ", f[x1, x2, x3] /. stationaryPoints[[i]]];
   Print["d^2L(x,λ)=", d2LMx // TraditionalForm];
   Print["PositiveDefinite:", PositiveDefiniteMatrixQ[d2LMx], "\n",
     "NegativeDefinite:", NegativeDefiniteMatrixQ[d2LMx]];)
 }
]
```

Рис. 4: Код: находим стационарные точки, подставляем их во второй дифференциал функции Лагранжа, упраздняя dx_3 с помощью уравнений связи, считаем значение целевой функции и проверяем положительную/отрицательную определённость матрицы второго дифференциала функции Лагранжа как квадратичной формы

Stationary point: $\{x_1 \to -0.965973,\ x_2 \to -0.198601,\ x_3 \to -0.165694,\ \lambda \to 4.03523\}$

Function value at this point: -5.36549

$$d^2L(x,\lambda) = \begin{pmatrix} 412.313 & 84.3444 \\ 84.3444 & 27.4114 \end{pmatrix}$$

PositiveDefinite:True
NegativeDefinite:False

Stationary point: $\{x_1 \to 0.16262,\ x_2 \to 0.46527,\ x_3 \to 0.870103,\ \lambda \to -3.14929\}$

Function value at this point: 4.64728

$$d^2L(x,\lambda) = \begin{pmatrix} -12.3789 & -0.22972 \\ -0.22972 & -4.95583 \end{pmatrix}$$

PositiveDefinite:False
NegativeDefinite:True

Stationary point: $\{x_1 \to 0.902445,\ x_2 \to -0.345794,\ x_3 \to -0.256944,\ \lambda \to 1.8919\}$

Function value at this point: -1.59219

$$d^2L(x,\lambda) = \begin{pmatrix} 93.8025 & -36.7919 \\ -36.7919 & 19.8815 \end{pmatrix}$$

PositiveDefinite:True
NegativeDefinite:False

Stationary point: $\{x_1 \to 0.360167,\ x_2 \to -0.817321,\ x_3 \to -0.44974,\ \lambda \to 0.223509\}$

Function value at this point: -1.1304

$$d^2L(x,\lambda) = \begin{pmatrix} -2.70095 & -6.47207 \\ -6.47207 & 17.134 \end{pmatrix}$$

PositiveDefinite:False
NegativeDefinite:False

Рис. 5: Вывод скрипта

Отсюда видим, что первая стационарная точка действительно является минимумом (что согласуется с решением, полученным специально предназначенной для этого функции), вторая — максимумом, третья — тоже минимумом, но локальным, а четвёртая — седлом. □