Методи оптимізації. Лекція 15 за 10.06.2022

Методи штрафних функцій

Здійснення мінімізації без обмежень є більш простою задачею, ніж мінімізація з обмеженнями. Тому природною є спроба зведення задачі умовної оптимізації до задачі безумовної оптимізації.

Методи штрафних функцій дозволяють звести задачу нелінійного програмування

$$f(x) \to min, \qquad (1)$$
$$g_i(x) \le 0,\ i = \overline{1,m}$$

до задачі або послідовності задач безумовної мінімізації певних допоміжних штрафних функцій. При цьому допоміжні функції підбираються так, що

1) на більшій частині допустимої множини задачі ці функції є близькими до нуля;

2) кожна з них достатньо швидко збільшується або при наближенні зсередини до границі допустимої множини (***внутрішні*** або ***бар'єрні штрафні функції***), або при виході за її границі (***зовнішні штрафні функції***);

3) ступінь близькості штрафу до нуля і швидкість його зростання залежать від значення штрафного параметру і збільшується із зростанням параметра.

У комбінованих методах штрафних функцій, які потрібно використовувати при обмеженнях-рівностях, у процесі мінімізації частина обмежень задовольняється, а частина – ні. Але при досягненні шуканого розв'язку всі умови в границях заданої точності задовольняються.

1. Метод зовнішньої точки

Основна ідея методу зовнішньої точки полягає в такому перетворенні цільової функції $f(x)$, при якому значення введеної перетвореної цільової функції в допустимій області точно або наближено дорівнюють значенням функції $f(x)$, у той час коли значення зовні допустимої області $X = \left\{ x \in E^n : g_i(x) \le 0, i = \overline{1,m} \right\}$ є дуже великими у порівнянні із значеннями функції $f(x)$. При виконанні цих умов мінімум вихідної функції $f(x)$ не буде істотно відрізнятися від мінімуму перетвореної функції.

Послідовність функцій $\left\{ P(x, r_k) \right\}$, $k = 1, 2, \ldots$, які визначені і невід'ємні на множині X, називається ***штрафною функцією*** множини X (***зовнішньою штрафною функцією***), якщо

$$\lim_{k \to \infty} P(x, r_k) = \begin{cases} 0, & \text{якщо } x \in X, \\ +\infty, & \text{якщо } x \notin X. \end{cases} \qquad (2)$$

Метод штрафних функцій розв'язання задачі (1) полягає в заміні цієї задачі послідовністю задач безумовної мінімізації функцій

$$f_k(x) = f(x) + P(x, r_k). \qquad (3)$$

При побудові допоміжних функцій бажано забезпечити потрібну гладкість, опуклість, зручність обчислення значень функції та потрібних похідних.

Якщо функції $g_i(x)$ не є опуклими, то штрафна функція також не буде опуклою на множині X. Тоді вона може мати локальні мінімуми, в той час, як вважається, що знаходиться глобальний мінімум задачі (3). Це порушує збіжність і є істотним недоліком методу штрафних функцій при застосуванні його до неопуклих задач.

У випадку, коли розглядається задача опуклого програмування, штрафна функція також буде опуклою. Але в цьому випадку виникає інша складність. Для отримання доброго наближення потрібно параметр r_k вибрати достатньо великим. При цьому всі похідні від $P(x, r_k)$ за змінною x також будуть великими, тому що вони пропорційні r_k. Але встановлено, що розмір околу, в якому збіжність надлінійна, є обернено пропорційним константі Ліпшиця других похідних, тобто в розглядуваному випадку цей окіл буде також малим, і навіть теоретично добре збіжний метод може стати неефективним.

Задача (3) розв'язується одним з методів безумовної оптимізації.

Якщо функція $f_k(x)$ є неперервно-диференційовною і обчислення її похідних не викликає істотних труднощів, то для безумовної мінімізації можна застосовувати методи першого порядку.

Якщо функція $f_k(x)$ є двічі неперервно-диференційовною, то можна застосувати метод Ньютона.

Негладкі функції штрафу призводять до необхідності мінімізації негладкої функції $f_k(x)$. У цьому випадку можна застосовувати методи нульового порядку (покоординатного спуску, конфігурацій, Розенброка та інші).

Теорема про збіжність. Нехай цільова функція $f(x)$ є неперервною, а штрафна функція $P(x, r_k)$ має такі властивості:

1) $P(x, r_k)$ неперервна за змінними x та r_k і $P(x, r_k) = 0$ при $x \in X$;

2) $P(x, r_k)$ монотонно зростає з зростанням r_k.

Нехай, окрім того, для довільної сталої C множина $X_C(k) = \left\{ x \in E^n : f_k(x) \le C \right\}$ є обмеженою. Тоді:

1) функція $f_k(x)$ досягає на E^n свого мінімального значення f_k^* в деякій точці x_k^*, при цьому $f_k^* \le f_*$, де $f_* = \min\limits_{x \in X} f(x)$ і $f_k^* \to f_*$;

2) будь-яка гранична точка x_* послідовності $\left\{ x_{r_k}^* \right\}$ при $r_k \to \infty$, $k \to \infty$ є точкою мінімуму x_* функції $f(x)$ множині X, вся послідовність $\left\{ x_{r_k}^* \right\}$ збігається до x_*.

Штрафну функцію $P(x,r_k)$ можна будувати різними способами. Наведемо ряд прикладів вибору $P(x,r_k)$ для задачі (1):

$$P(x,r_k) = r_k \sum_{i=1}^{m} \left[max\left\{ g_i(x), 0 \right\} \right]^q,\ q > 0,$$

$$P(x,r_k) = \frac{1}{r_k} exp\left[r_k \sum_{i=1}^{m} max\left\{ g_i(x), 0 \right\} \right],$$

$$P(x,r_k) = r_k \max_{1 \le i \le m} \left\{ max\left\{ g_i(x), 0 \right\} \right\}.$$

Якщо задача (1) містить обмеження-рівності, то штрафну функцію можна вибрати у вигляді:

$$P(x,r_k) = r_k \sum_{i=1}^{m} \left(g_i(x) \right)^p, \text{ де } p \ge 1 \text{ – фіксоване число.}$$

Алгоритм методу зовнішньої точки

Нехай функції $f(x)$, $g_i(x)$, $i = \overline{1,m}$ є неперервними.

Початковий етап. Задати початкову точку $x^{(1)} \in E^n$, штрафний параметр $r_1 > 0$ і число $\beta > 0$. Вибрати $\varepsilon > 0$ як константу зупинки. Покласти $k = 1$ та перейти до основного етапу.

Основний етап

Крок 1. При початковій точці $x^{(k)}$ розв'язати таку задачу безумовної оптимізації:

$$f_k(x) = f(x) + P(x,r_k) \to min.$$

Покласти $x^{(k+1)}$ таким, що дорівнює оптимальному розв'язку цієї задачі і перейти до кроку 2.

Крок 2. Якщо $P(x,r_k) \le \varepsilon$, то зупинитися, у протилежному випадку покласти $r_{k+1} = \beta r_k$, замінити k на $k+1$та перейти до кроку 1.

Алгоритм описаний.

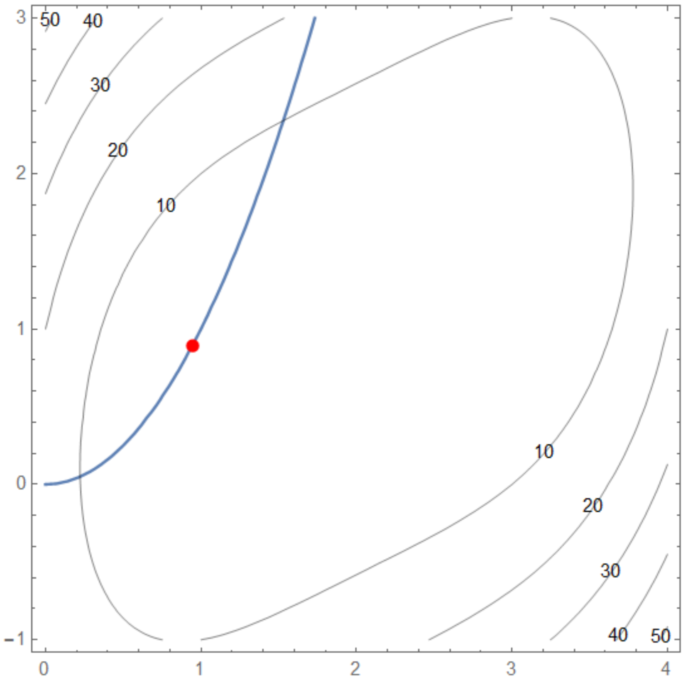
Приклад. Розглянемо таку задачу:

$$f(x) = (x_1 - 2)^4 + (x_1 - 2x_2)^2 \to min$$

за умови

$$x_1^2 - x_2 = 0,\ x = (x_1, x_2) \in E^2.$$

На рисунку наведено геометричну інтерпретацію задачі: лінії рівня цільової функції $(x_1 - 2)^4 + (x_1 - 2x_2)^2$, графік кривої $x_2 = x_1^2$. Червоною точкою позначено оптимальний розв'язок задачі $x^* = (0.9456,\ 0.8941)$, $f(x^*) = 1.9462$.



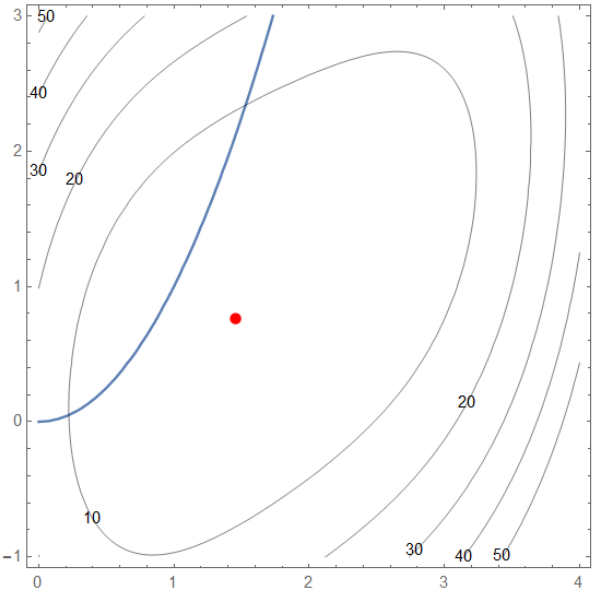
За методом зовнішньої точки на k-тій ітерації при заданому значенні параметра штрафу r_k буде розв'язуватися така задача:

$$(x_1 - 2)^4 + (x_1 - 2x_2)^2 + r_k\left(x_1^2 - x_2\right)^2 \to min\,.$$

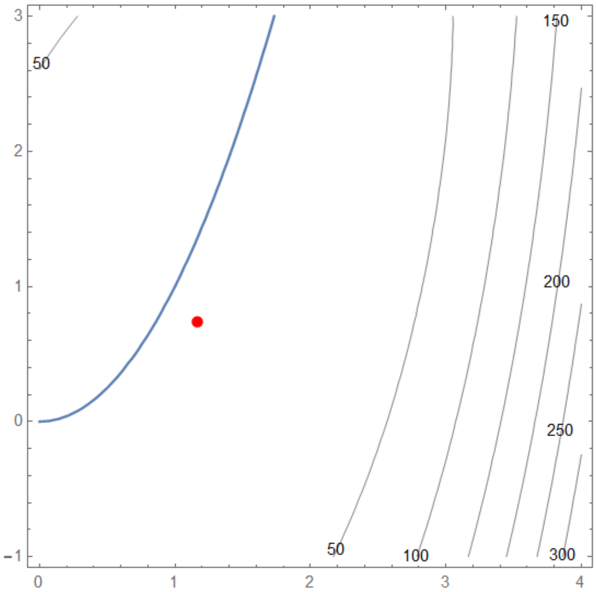
У таблиці наведені результати обчислень за методом зовнішньої точки з початкової точки $x^{(1)} = (0,\,0)$, число $\beta = 10$. Задача безумовної оптимізації розв'язувалась градієнтним методом з дробленням кроку, точність обчислень $\varepsilon = 10^{-4}$.

Таблиця

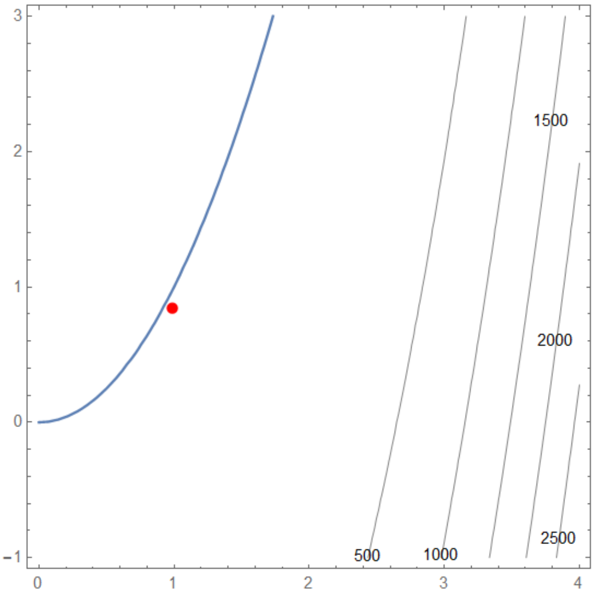
k	r_k	$x^{(k+1)}$	$f\left(x^{(k+1)}\right)$	$r_k\ P\left(x^{(k+1)}, r_k\right)$
1	0,1	(1.4538, 0.7607)	0.0936	0.1830
2	1,0	(1.1687, 0.7407)	0.5754	0.3908
3	10,0	(0.9905,0.8422)	1.5202	0.1928
4	100,0	(0.9488,0.8834)	1.8914	0.0270
5	1000,0	(0.9461,0.8934)	1.9404	0.0029



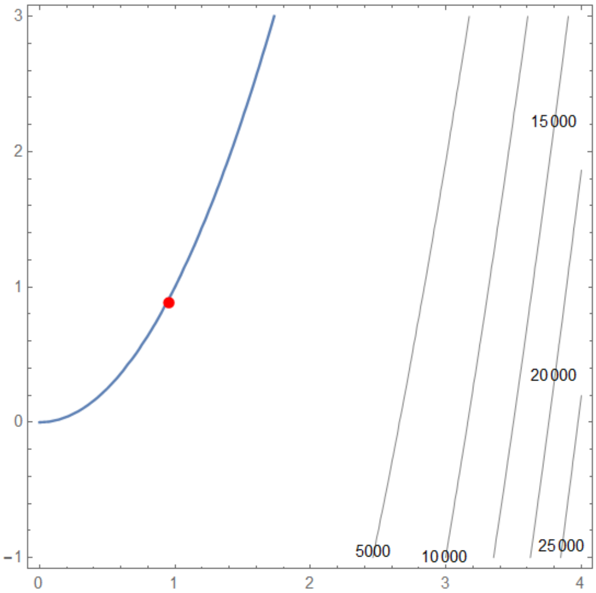
$r_1 = 0.1$



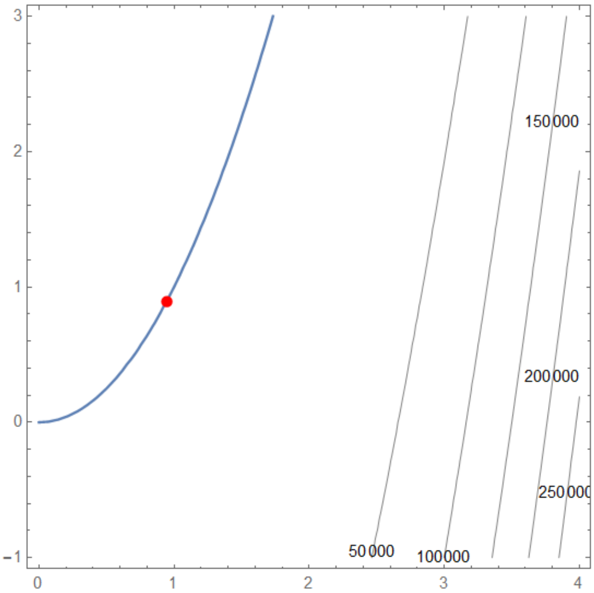
$r_2 = 1$



$r_3 = 10$



$r_4 = 100$



$r_5 = 1000$

2. Метод внутрішньої точки (метод бар'єрів)

Головна незручність попереднього методу полягає в тому, що точка мінімуму x_* функції $f(x)$ на множині X апроксимується зовні, тобто різні проміжні значення $x_1^*,\dots,\ x_k^*$, які отримано при коефіцієнтах штрафу $r_1,\dots,\ r_k$, не належать допустимій множині X. Це і привело до необхідності розробки інших методів штрафу, в яких оптимум апроксимується зсередини.

Функцію $B(x,r_k)$ назвемо ***функцією бар'єру***, якщо

1) $B(x,r_k)$ визначена і невід'ємна в усіх внутрішніх точках допустимої множини,

2) $B(x,r_k)\to+\infty$, якщо x прямує до границі допустимої множини $x\in E^n$.

При неперервних функціях $g_i(x)\le 0,\quad i=\overline{1,m}$ функція $B(x,r_k)$ неперервна у внутрішніх точках множини X.

Прикладами функцій бар'єру є такі функції:

$$B(x,r_k)=\frac{1}{r_k}\sum_{i=1}^{m}\frac{1}{-g_i(x)},$$

$$B(x,r_k)=\frac{1}{r_k}\sum_{i=1}^{m}\frac{1}{g_i^2(x)},$$

$$B(x,r_k)=-\frac{1}{r_k}\sum_{i=1}^{m}ln\left[-g_i(x)\right].$$

На відміну від методу зовнішньої точки за початкову точку $x^{(0)}$ вибирається ***допустима*** точка, тобто $x^{(0)}\in X$. Мінімум функції $f_k(x)$ для методу внутрішньої точки досягається зсередини допустимої області, але не лежить на її границі при усіх $r_k>0$. Істотною відмінністю від методу зовнішньої точки є ***вимога існування внутрішніх точок допустимої області***.

Алгоритм методу внутрішньої точки співпадає з алгоритмом методу зовнішньої за винятком того, що в алгоритмі методу внутрішньої точки $x^{(0)}\in X$.

Пошук початкового наближення $x^{(0)}$, тобто точки $x^{(0)}\in X$, такої що $g_i\left(x^{(0)}\right)<0$, можна здійснити за таким алгоритмом.

Алгоритм вибору початкової точки в методі внутрішньої точки

Початковий етап. Задати точку $x^{(1)}$, покласти $k=1$ та перейти до основного етапу.

Основний етап

Крок 1. Покласти $I=\left\{ i : g_i\left(x^{(k)}\right)<0 \right\}$, Якщо $I=\left\{ 1,\ldots,m \right\}$, то зупинитися: точка $x^{(k)}$ є шуканим початковим наближенням. У протилежному випадку вибрати $j \notin I$ і перейти до кроку 2.

Крок 2. Використати метод бар'єрів для розв'язання такої задачі при початковій точці $x^{(k)}$:

$$g_j\left(x\right)\to min, \tag{4}$$
$$g_i\left(x\right)<0,\ i\in I .$$

Покласти $x^{(k+1)}$ таким, що дорівнює оптимальному розв'язку задачі (4). Якщо $g_j\left(x^{(k+1)}\right)\geq 0$, то зупинитися, тому що множина $\left\{x\in E^n : g_i\left(x\right)<0, i=\overline{1,m}\right\}$ є порожньою. У протилежному випадку замінити k на $k+1$та перейти до кроку 1.

Алгоритм описаний.

Алгоритм зупиниться не більше, ніж через m ітерацій або знайшовши допустиму початкову точку, яка задовольняє нерівності $g_i\left(x^{(0)}\right)<0$, або встановивши факт відсутності таких точок.

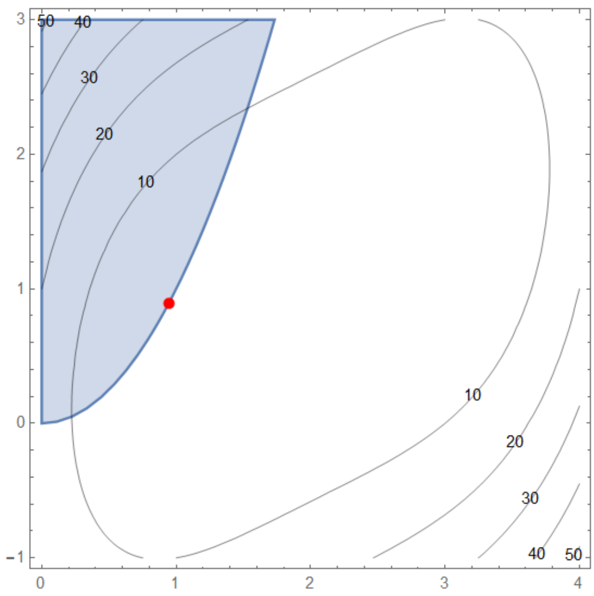
Приклад. Розглянемо таку задачу:

$$f\left(x\right)=\left(x_1-2\right)^4+\left(x_1-2x_2\right)^2\to min$$

за умови

$$x_1^{\,2}-x_2\leq 0,\ x=\left(x_1,x_2\right)\in E^2 .$$

На рисунку наведено геометричну інтерпретацію задачі: лінії рівня цільової функції $\left(x_1-2\right)^4+\left(x_1-2x_2\right)^2$, графік кривої $x_2=x_1^{\,2}$. Червоною точкою позначено оптимальний розв'язок задачі $x^*\approx\left(0.9456,\ 0.8941\right)$, $f\left(x^*\right)\approx 1.9462$.



За методом внутрішньої точки на k-тій ітерації при заданому значенні параметра штрафу r_k буде розв'язуватися така задача:

$$\left(x_1-2\right)^4+\left(x_1-2x_2\right)^2-\frac{1}{r_k\left(x_1^2-x_2\right)}\to min.$$

У таблиці наведені результати обчислень за методом внутрішньої точки з початкової точки $x^{(1)}=\left(0,1\right)$.

Таблиця

k	r_k	$x^{(k+1)}$	$f\left(x^{(k+1)}\right)$	$B\left(x^{(k+1)},r_k\right)$
1	0,1	(0.7079, 1.5315)	18.039	9.7058
2	1,0	(0.8282, 1.1098)	6.1806	2.3592
3	3,0	(0.8876, 1.0037)	3.9340	1.0638
4	5,0	(0.8828, 0.9923)	3.7110	0.9389
5	7,0	(0.8911, 0.9773)	3.4227	0.7796

3. Метод послідовної безумовної оптимізації

(метод Фіако і Мак-Кормика)

Методи зовнішньої і внутрішньої точок базуються на істотно різних принципах. У методах внутрішньої точки штрафний член перешкоджає порушенню обмежень, тому все наближення до оптимуму знаходяться всередині допустимої області. У методах зовнішньої точки, навпаки, в процесі розв'язання утворюється послідовність точок, яка виходить за межі допустимої області, штрафний член в цьому випадку запобігає блуканню точок занадто далеко від допустимої області. Крім того, в методах внутрішньої точки за початкове наближення для знаходження мінімуму розширеної функції необхідно вибирати допустиму точку області, в той час як в методах зовнішньої точки взагалі не потрібно вирішення питання про допустимість точок послідовності.

Зазначимо, що для задач з обмеженнями у формі рівностей методи внутрішньої точки не можуть бути прийнятні, так як для застосування цих методів потрібно існування внутрішності допустимої області. Тому для розв'язання задач математичного програмування, що включають в себе як обмеження-нерівності, так і обмеження-рівністі, краще застосовують комбіновані методи.

Розглянемо задачу:

$$f\left(x\right)\to min, \tag{5}$$

$$g_i\left(x\right)\le 0,\ i=\overline{1,p},$$

$$h_i\left(x\right)=0,\ i=\overline{p+1,m}.$$

Тут $h_i\left(x\right),\ i=\overline{p+1,m}$ – ***лінійні*** функції.

Розширеними штрафними функціями в комбінованих методах можуть бути такі функції:

$$f_k(x) = f(x) + \frac{1}{r_k}\sum_{i=1}^{p}\frac{1}{-g_i(x)} + \sqrt{r_k}\sum_{i=p+1}^{m} h_i^2(x),$$

$$f_k(x) = f(x) - \frac{1}{r_k}\sum_{i=1}^{p} ln\left[-g_i(x)\right] + \sqrt{r_k}\sum_{i=p+1}^{m} h_i^2(x).$$

Процедура мінімізації функції $f_k(x)$ починається з внутрішньої початкової точки $x^{(0)}$, в якій задовольняються всі обмеження у вигляді нерівностей.

Швидкість збіжності методу залежить від початкового вибору параметру r_0 та способу зміни параметру r_k.

Існують різні способи вибору початкового параметру r_0. Наприклад,

1. $r_0 = 1$.
2. $$r_0 = \frac{f'\left(x^{(0)}\right)^T \cdot R'\left(x^{(0)}\right)}{\left\| R'\left(x^{(0)}\right)\right\|^2},$$

де $R\left(x^{(0)}\right) = \sum_{i=1}^{p}\frac{1}{-g_i\left(x^{(0)}\right)}$.

3. $$r_0 = \left[\frac{f'\left(x^{(0)}\right)^T \cdot \left[R''\left(x^{(0)}\right)\right]^{-1} f'\left(x^{(0)}\right)}{R'\left(x^{(0)}\right)^T \left[R''\left(x^{(0)}\right)\right]^{-1} R'\left(x^{(0)}\right)}\right]^{-1}.$$

Що стосується подальшого вибору значень параметра r_k, то як показала практика застосування методів штрафних функцій, ці методи є досить ефективними, коли послідовність r_1, r_2, …, r_k визначається простим співвідношенням $r_{k+1} = \beta r_k$, де $\beta > 1$. На практиці часто $\beta = 4$.

Умови збіжності методу послідовної безумовної мінімізації

Нехай функції $f(x)$, $g_i(x)$, $i = \overline{1,p}$, $h_i(x)$, $i = \overline{p+1,m}$ є двічі неперервно-диференційовними функціями і виконані такі умови:

1) множина X має внутрішні точки;
2) для будь-якого скінченного C і будь-якого $r_k > 0$ множина точок $X_C(k) = \left\{ x \in E^n : g_i(x) \le 0, i = \overline{1,p},\ f(x) + \sqrt{r_k}\sum_{i=p+1}^{m} h_i^2(x) \le C \right\}$ є обмеженою;
3) функція $f(x)$ є опуклою і сума $\sum_{i=p+1}^{m} h_i^2(x)$ також є опуклою;
4) функції $g_i(x)$, $i = \overline{1,m}$ є опуклими;

5) матриця Гессе розширеної функції $f_k\left(x\right)$ не обертається в нуль для жодної точки, яка належить множині $X_1 = \left\{ x \in E^n : g_i\left(x\right) \le 0, i = \overline{1,p} \right\}$.

Теорема. Якщо задача нелінійного програмування задовольняє умовам 1–5, то

1) кожна розширена функція $f_k\left(x\right)$ має мінімум у деякій точці x_k^* множини X;

2) якщо $\left\{ r_k \right\}$ – строго зростаюча послідовність, тоді

$$\lim_{k \to \infty} f_k\left(x_k^*\right) = \min_{x \in X} f\left(x\right) = f\left(x_*\right) = f_*,$$

тобто послідовність безумовних мінімумів функції $f_k\left(x\right)$ збігається до умовного мінімуму функції $f\left(x\right)$.

Критерієм закінчення ітераційного процесу може бути такий критерій:

$$\frac{1}{r_k} \sum_{i=1}^{p} \frac{1}{g_i\left(x^{(k)}\right)} \le \varepsilon,$$

де ε – задана константа.