Многомерная условная оптимизация

Найти минимум критерия оптимальности Ф(X), определенного на множестве D евклидова пространства Rn

$$\min_{\overline{X}\in D} \Phi\left(\overline{X}\right)=\Phi\left(\overline{X}*\right)=\Phi*$$

где область определения

$$D=\left\{\overline{X}\mid h_i\left(\overline{X}\right)=0,\ g_j\left(\overline{X}\right)\geq 0,\ i\in[1,\ldots,m],\ j\in[1,\ldots,l]\right\}$$

1 Метод неопределенных множителей Лагранжа

1.1 Рассмотрим задачу оптимизации с ограничениями типа равенств.

$$\min_{\overline{X}\in D} \Phi\left(\overline{X}\right)=\Phi\left(\overline{X}*\right)=\Phi*$$

где область определения

$$D=\left\{\overline{X}\mid h_i\left(\overline{X}\right)=0,\ i\in[1,\ldots,m]\right\}$$

Множители Лагранжа и функции Лагранжа

Функция Лагранжа для задачи ограничениями типа равенств определяется формулой:

$$L\left(\overline{X},\overline{\lambda}\right)=\Phi\left(\overline{X}\right)+\sum_{j=1}^{m}\lambda_j h_j\left(\overline{X}\right)=\Phi\left(\overline{X}\right)+\overline{\lambda}^T\overline{h}\left(\overline{X}\right)$$

где $\overline{\lambda}=\left\{\lambda_j,\ j=1,\ldots,m\right\}$ – вектор множителей Лагранжа.

Условие регулярности ограничивающих функций

Если точка X*∈D, то условие линейной независимости векторов

$$\nabla h_j\left(X*\right),\ j=1,\ldots,m$$

называется условием регулярности задачи в точке X*.

Данное условие означает, в частности, что количество ограничивающих функций, проходящих через точку X* , не может быть больше размерности вектора варьируемых параметров, т.е. должно быть выполнено неравенство m≤n.

См рис.а - количество ограничивающих функций, проходящих через точку X*, превышает размерность вектора варьируемых параметров, рис.б - в точке X* градиенты $\nabla h_1\left(X*\right)$ и $\nabla h_2\left(X*\right)$ограничивающих функций коллинеарны

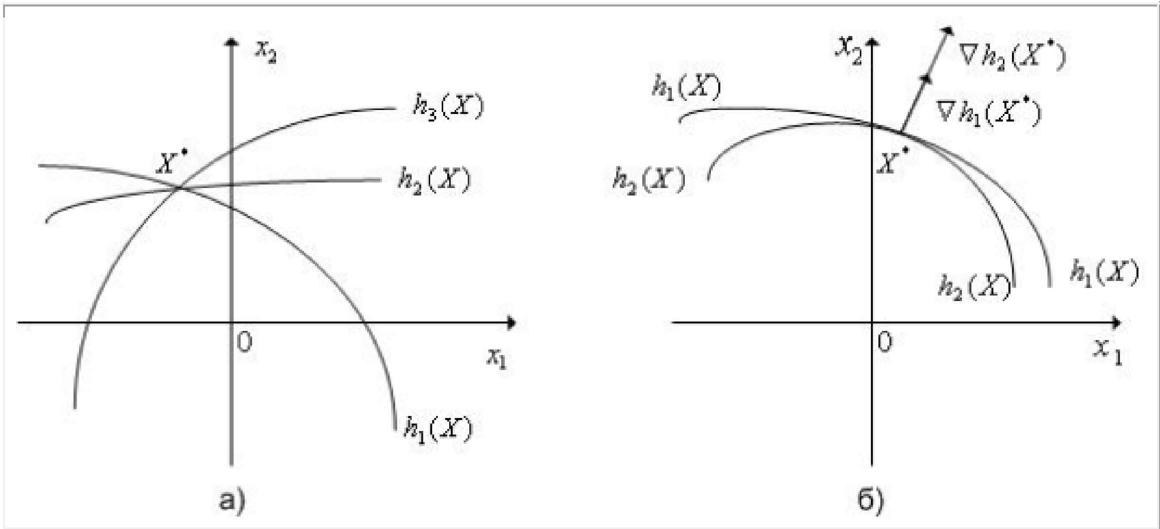


Рис. Пример невыполнения условий регулярности системы функций h(X) в точке X*

Важное место в теории и практике решения задач нелинейного программирования с ограничениями типа равенств занимает следующая теорема (правило Лагранжа для задачи оптимизации с ограничениями типа равенств).

Теорема (правило множителей Лагранжа)

Пусть функция Ф(X) и функции $h_j(X^*), j=1,\ldots,m$ имеют непрерывные частные производные в некоторой окрестности точки X* и пусть эта точка является точкой локального минимума функции Ф(X) при условии h(X*)=0 . Пусть, кроме того, выполняется условие регулярности системы функций h(X) в точке X*.

Тогда существуют такие множители Лагранжа $\lambda_j,\ j=1,\ldots,m$, *не все из которых равные нулю одновременно,* что для функции Лагранжа L(X, λ) точка X* является стационарной точкой функции, т.е.

$$\nabla_X L(X^*,\lambda) = \nabla \Phi(X^*) + \sum_{j=1}^{m} \lambda_j \nabla h_j(X^*) = 0$$

Отметим, что теорема *не требует знакоопределенности* (т.е. положительности или отрицательности) множителей Лагранжа $\lambda_j,\ j=1,\ldots,m$. Теорема требуется лишь того, чтобы не все из этих множителей равнялись нулю одновременно.

Пример

Рассмотрим в качестве минимизируемой функции Ф(X) функцию Розенброка:

$$\Phi(x_1,x_2) = 100\left(x_2 - x_1^2\right)^2 + (1 - x_1)^2$$

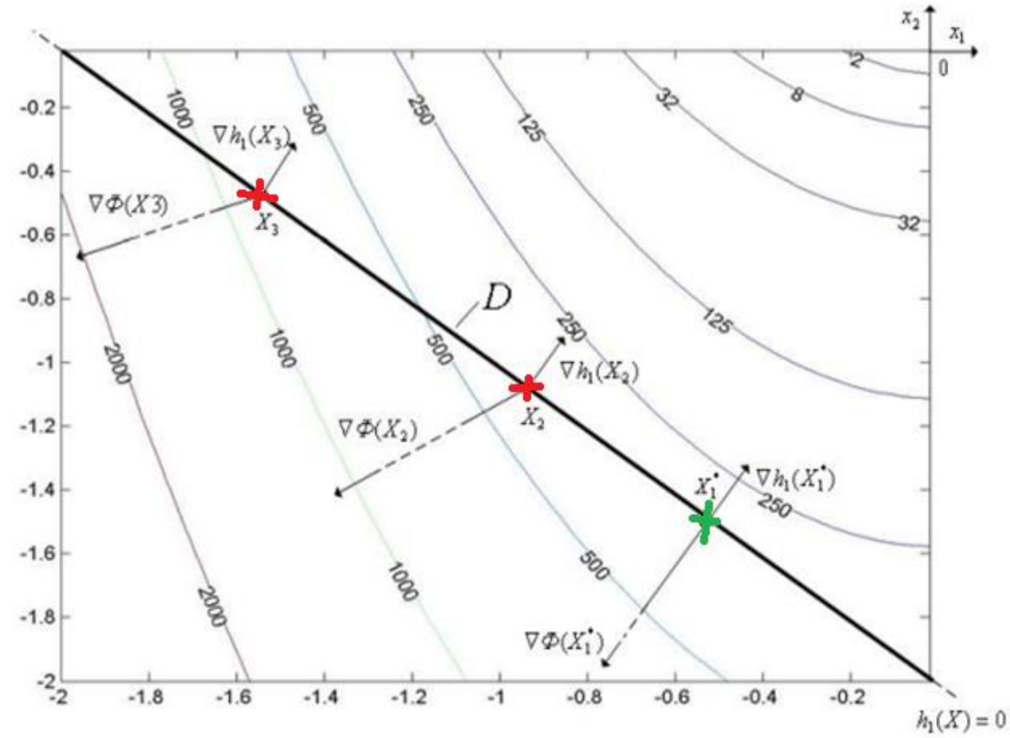
Положим, что имеется только одно ограничение типа равенств, которое задается с помощью функции

$$h_1(X) = x_1 + x_2 + 2 = 0$$

Легко видеть, что градиенты функций Φ и h равны, соответственно:

$$\nabla \Phi(X) = \begin{pmatrix} -400(x_2 - x_1)x_1 - 2(1 - x_1) \\ 200(x_2 - x_1^{\ 2}) \end{pmatrix}$$

$$\nabla h_1(X) = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$



В точках X2 и X3 векторы градиента функций Ф(X) и h(X) не коллинеарны. Поэтому для этих точек не существует не равный нулю множитель Лагранжа λ_1 , при котором функция Лагранжа равна нулю:

$$\nabla \Phi(X) + \lambda_1 \nabla h_1(X) = 0$$

Поэтому точки X2 и X3 не могут быть точками локального минимума для рассматриваемой задачи.

В точке X_1* векторы градиента функций Ф(X) и h(X) коллинеарны и поэтому существует не равный нулю множитель λ_1, при котором справедливо равенство

$$\nabla \Phi(X^*) + \lambda_1 \nabla h_1(X^*) = 0$$

Отметим, что, например, в точке (-1,-1) , градиент функции Розенброка равен (-402, -400).

Теорема (правило множителей Лагранжа) означает, что в ее условиях вместо задачи условной оптимизации можно решать задачу безусловной оптимизации

$$\min_{\overline{X} \in R^n} \left(\Phi(\overline{X}) + \sum_{j=1}^{m} \lambda_j h_j(X) \right)$$

Необходимым условием существования локального минимума этой задачи в некоторой точке X*∈Rn является условие

$$\nabla_X \left(\Phi\left(\overline{X}\right) + \sum_{j=1}^{m} \lambda_j h_j\left(X\right) \right) = 0$$

Широко известна другая форма теоремы, которую мы сформулируем в виде следствия этой теоремы.

Следствие

В условиях теоремы (правила множителей Лагранжа) существуют такие множители Лагранжа $\lambda_j, j = 1,\ldots,m$, *не все из которых равные нулю одновременно*, что имеют место следующие равенства:

$$\nabla_X L\left(X^*, \lambda\right) = 0$$

$$\nabla_\lambda L\left(X^*, \lambda\right) = h\left(X^*\right) = 0$$

Здесь первое равенство повторяет условие стационарности L (см правило), а справедливость второго равенства следует из того факта, что по условиям теоремы точка X* удовлетворяет всем ограничениям, т.е. $h_j\left(\overline{X}^*\right) = 0,\ j = 1,\ldots,m$

Заметим, что из второго равенства следует справедливость еще одного полезного равенства

$$\lambda^T \nabla_\lambda L\left(\overline{X}^*, \lambda\right) = \sum_{j=1}^{m} \lambda_j h_j\left(\overline{X}^*\right) = \overline{\lambda}^T \overline{h}\left(\overline{X}^*\right) = 0$$

1.2 Задача нелинейного программирования с ограничениями типа неравенств

Рассмотрим задачу нелинейного программирования:

$$\min_{\overline{X} \in D} \Phi\left(\overline{X}\right) = \Phi\left(\overline{X}^*\right) = \Phi^*$$

где область определения

$D = \left\{ \overline{X} \mid g_j\left(\overline{X}\right) \geq 0,\ j \in [1,\ldots,m] \right\}$ – непустое ограниченное замкнутое множество.

Множители Лагранжа и функции Лагранжа для задачи нелинейного программирования с ограничениями типа неравенств

Функция Лагранжа для задачи с ограничениями определяется формулой:

$$L\left(X, \lambda\right) = \Phi\left(X\right) + \sum_{j=1}^{m} \lambda_j g_j\left(X\right) = \Phi\left(X\right) + \overline{\lambda}^T \overline{g}\left(X\right)$$

Где $\overline{\lambda} = \left\{ \lambda_j,\ j = 1,\ldots,m \right\}$ – вектор множителей Лагранжа.

Активные и неактивные ограничения

В точке локального минимума каждое из ограничений $g_j(\overline{X}) \geq 0$ выполняется либо в виде равенства $g_j(\overline{X}) = 0$, либо в виде неравенства $g_j(\overline{X}) > 0$

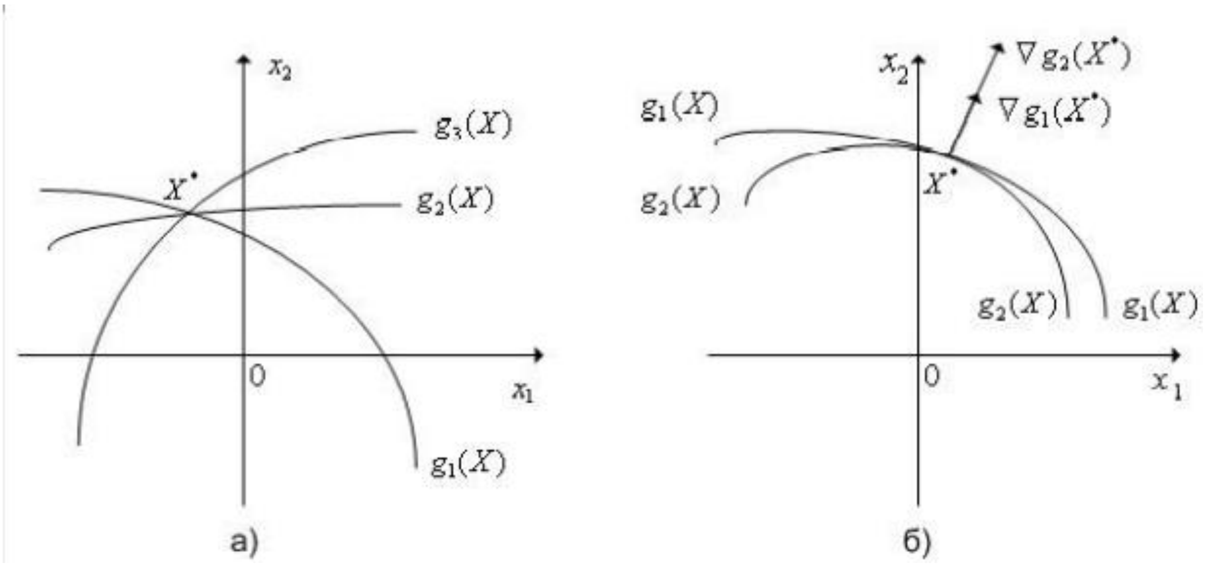
Ограничения первого вида (=) называются ***активными ограничениями***.
Остальные ограничения называются ***неактивными ограничениями***.

Условие регулярности для задачи нелинейного программирования с ограничениями типа неравенств

Если точка X∗∈D и s ограничений $g_j(\overline{X}) \geq 0$ из m активны, то условие линейной независимости векторов называется условием регулярности ограничивающих функций $g_j(X^*) \geq 0,\ j = 1,\ldots,s$ в точке X∗.

Это условие означает, что, например, при n=2 количество ограничивающих функций, проходящих через точку X∗, не должно превышать 2 и в точке X∗ векторы $\nabla g_1(X^*)$ и $\nabla g_2(X^*)$ не должны быть коллинеарны.

Например, на рис. в ситуации (а) количество ограничивающих функций, проходящих через точку X∗ , превышает размерность вектора варьируемых параметров, в ситуации (б) в точке X∗ градиенты $\nabla g_1(X^*)$, $\nabla g_2(X^*)$ ограничивающих функций коллинеарны.



Ситуации, в которых не выполняется условие регулярности двумерной задачи

Большое значение в теории и практике решения задач нелинейного программирования имеет следующая теорема.

Теорема (Куна-Таккера для задачи условной оптимизации с ограничениями типа неравенств)

Пусть функция Ф(X) и функции $g_j(\overline{X}) \geq 0,\ j = 1,\ldots,m$ имеют непрерывные

частные производные в некоторой окрестности точки X∗ и пусть эта точка является точкой локального минимума функции Ф(X) при ограничениях $g_j\left(\overline{X}\right) \geq 0$, удовлетворяющих в точке X* условию регулярности ограничивающих функций. Тогда существуют такие неотрицательные множители Лагранжа λi, i=1,…,m, что для функции Лагранжа L(X,λ) точка X* является стационарной точкой функции, т.е.

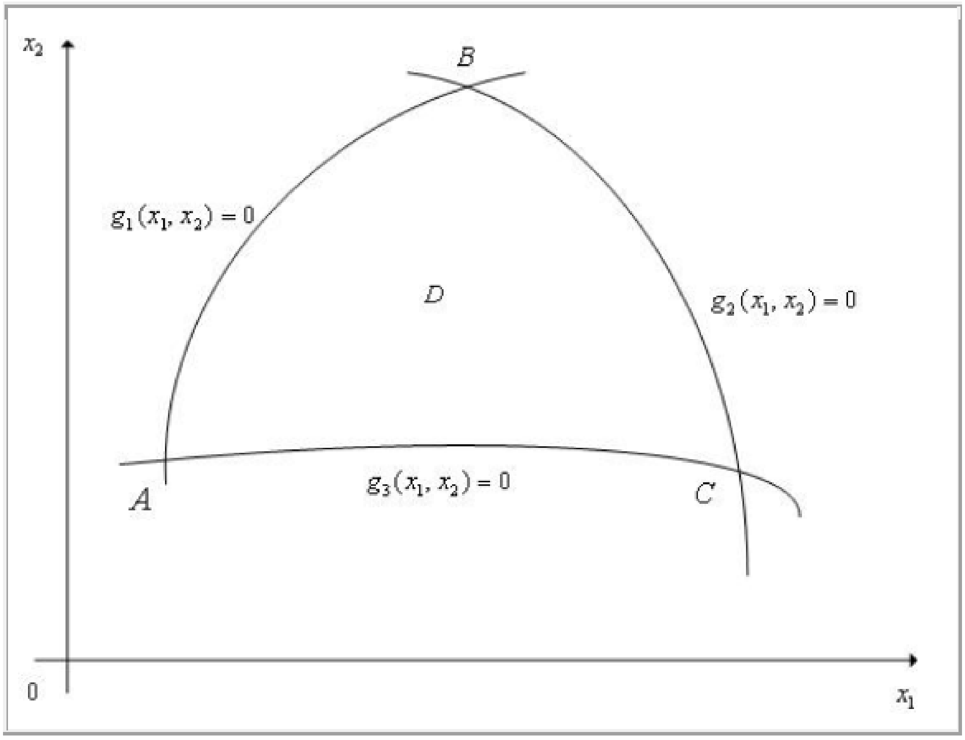
$$\nabla_X L\left(X^*,\lambda\right) = \nabla \Phi\left(X^*\right) + \sum_{j=1}^{m} \lambda_j \nabla g_j\left(X^*\right) = 0$$

Заметим, что в отличие от правила множителей Лагранжа, теорема Куна-Таккера *требует знакоопределенности множителей Лагранжа* λi, i=1,…,m.

Отметим также, что *теорема не запрещает того, чтобы все множители Лагранжа были равны нулю.*

Пример

Рассмотрим двумерную (n=2) задачу нелинейного программирования в которой область допустимых значений D задается тремя ограничивающими функциями g, а множество D имеет вид, представленный на рис.



Для всех граничных точек области D, очевидно, выполняются условия регулярности ограничивающих функций.

Если точка X* находится внутри множества D (т.е. является стационарной точкой функции Ф(X)), то теорема будет справедлива, если положить все множители Лагранжа равными нулю.

Если точка X* находится на одной из дуг, например, на дуге AB, т.е. пусть ограничение g1≥ 0 является активным ограничением, а остальные ограничения – неактивными ограничениями.
Тогда в этой точке g1(X*)=0 и справедливость теоремы вытекает из правила множителей Лагранжа для задачи с ограничениями типа равенств, если положить λ2=λ3=0.

Если точка X* находится в одной из угловых точек множества D, например, в точке B, т.е. g1(X)≥ 0 и g2(X)≥ 0 являются активными ограничениями, а ограничение g3(X)≥ 0 – неактивным ограничением. Тогда можно положить λ3=0 и справедливость теоремы вытекает из правила множителей Лагранжа для задачи с ограничениями типа равенств.

Теорема Куна-Таккера означает, что в ее условиях вместо задачи условной оптимизации можно решать задачу безусловной оптимизации

$$\min_{\bar{X}\in R^n}\left(\Phi(\bar{X})+\sum_{j=1}^{m}\lambda_j g_j(X)\right)$$

Необходимым условием существования локального минимума этой задачи в некоторой точке X*∈Rn является условие

$$\nabla_X\left(\Phi(X^*)+\sum_{j=1}^{m}\lambda_j g_j(X^*)\right)=0$$

Широко известна другая форма теоремы (см. Следствие)
Следствие. В условиях теоремы К-Т существуют такие неотрицательные множители Лагранжа λi, i=1,…,m, что имеют место следующие равенства:

$$\nabla_X L(X^*,\lambda)=0$$

$$\nabla_\lambda L(X^*,\lambda)=g(X^*)\geq 0$$

Здесь первое равенство повторяется, а справедливость равенства следует из того, что по условиям теоремы точка X* удовлетворяет всем ограничениям, т.е. gi(X)≥ 0, i=1…m.

Заметим, что из теоремы К-Т следует справедливость еще одного полезного равенства

$$\lambda^T\nabla_\lambda L(X^*,\lambda)=\sum_{j=1}^{m}\lambda_j g_j(X^*)=\lambda^T g(X^*)\geq 0$$

1.3 Теорема Куна-Таккера для общей задачи нелинейного программирования

Рассмотрим общую задачу нелинейного программирования:

$$\min_{\bar{X}\in D} \Phi(\bar{X}) = \Phi(\bar{X}^*) = \Phi^*$$

где Ф(X) – произвольная функция,

$$D = \{\bar{X} \mid h_i(\bar{X}) = 0,\ g_j(\bar{X}) \ge 0,\ i \in [1,\ldots,m],\ j \in [1,\ldots,l]\}$$

непустое ограниченное замкнутое множество

Множители Лагранжа и функции Лагранжа для общей задачи нелинейного программирования

Функция Лагранжа для задачи с ограничениями D определяется формулой

$$L(X,\lambda,\mu) = \Phi(X) + \sum_{i=1}^{m} \lambda_i g_i(X) + \sum_{j=1}^{l} \mu_j h_j(X) = \Phi(X) + \bar{\lambda}^T \bar{g}(X) + \bar{\mu}^T \bar{h}(X)$$

Где $\bar{\lambda} = (\lambda_i, i = 1,\ldots,m),\ \bar{\mu} = (\mu_j,\ j = 1,\ldots,l)$

– векторы множителей Лагранжа, соответственно.

Условия регулярности для общей задачи нелинейного программирования

Если точка X∗∈D и ограничения gj(X*) , j=1…s, являются активными ограничениями, то условие линейной независимости векторов $\nabla g_i(X^*), i = 1,\ldots, s$

а также условие линейной независимости векторов $\nabla h_j(X^*),\ j = 1,\ldots,l$

называются *условиями регулярности ограничивающих функций* в точке X∗.

Теорема (Куна-Таккера)

Пусть функции Ф(X) и функции $g_i(\bar{X}) \ge 0,\ h_j(\bar{X}) = 0,\ i = 1,\ldots,m,\ j = 1,\ldots,l$

имеют непрерывные частные производные в некоторой окрестности точки X∗∈D и пусть эта точка является точкой локального минимума функции Ф(X).

Пусть, кроме того, выполняются условия регулярности ограничивающих функций g и h в точке X∗.

Тогда существуют такие множители Лагранжа λi и μj, не все из которых равные нулю одновременно, что для функции Лагранжа L(X, λ,μ) точка X∗ является стационарной точкой функции, т.е.

$$\nabla_X L(X,\lambda,\mu) = \nabla\Phi(X) + \sum_{i=1}^{m} \lambda_i \nabla g_i(X) + \sum_{j=1}^{l} \mu_j \nabla h_j(X) = 0$$

Теорема К-Т означает, что в ее условиях ***вместо задачи условной оптимизации можно решать задачу безусловной оптимизации***

$$\min_{\overline{X}\in R^n}\left(\Phi\left(\overline{X}\right)+\sum_{i=1}^{m}\lambda_i g_i\left(X\right)+\sum_{j=1}^{l}\mu_j h_j\left(X\right)\right)$$

Необходимым условием существования локального минимума этой задачи в некоторой точке $X*\in D$ является условие

$$\nabla\Phi\left(\overline{X}\right)+\sum_{i=1}^{m}\lambda_i\nabla g_i\left(X\right)+\sum_{j=1}^{l}\mu_j\nabla h_j\left(X\right)=0$$

Схема решения задачи нелинейного программирования с ограничениями типа неравенств с использованием теоремы Куна-Таккера:

1. Записываем функцию Лагранжа $L(X, \lambda)$
2. Находим градиенты Ф и g
3. Находим стационарные точки функции Лагранжа, т.е. точки, в которых градиент функции Лагранжа равен нулю.
4. Находим точки, в которых нарушаются условия регулярности ограничивающих функций.
5. Во всех стационарных точках функции, а также точках нарушения условий регулярности ограничивающих функций вычисляем значения функции Ф(X) и выбираем ту (или те), в которой значение функции наименьшее.

Задача

Пусть дана двумерная задача нелинейного программирования с ограничениями типа неравенств - найти

$$\min_{\overline{X}\in D}\Phi\left(\overline{X}\right)=\Phi\left(\overline{X}*\right)=\Phi*$$

где

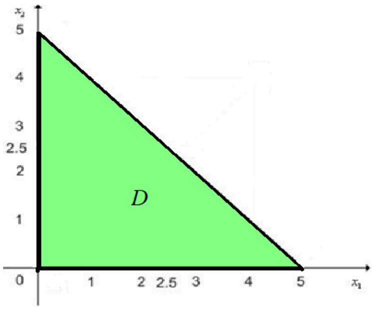
$$\Phi\left(\overline{X}\right)=\Phi\left(x_1,x_2\right)=\left(x_1-4\right)^2+\left(x_2-4\right)^2$$

где область определения

$$D=\left\{\overline{X}\mid g_1\left(\overline{X}\right)=x_1\geq 0,\ g_2\left(\overline{X}\right)=x_2\geq 0, g_3\left(\overline{X}\right)=-x_1-x_2+5\geq 0\right\}$$

Задача является задачей выпуклого программирования: множество D есть выпуклый многогранник, а функция Ф(X) – квадратичная.

Область допустимых значений вектора варьируемых параметров, формируемая ограничениями, имеет вид, представленный на рис.



A(4,4) – точка минимума Ф(Х) без учета ограничений.

Записываем функцию Лагранжа:

$$L(\bar{X},\lambda) = (x_1-4)^2 + (x_2-4)^2 + \lambda_1 x_1 + \lambda_2 x_2 + \lambda_3(-x_1 - x_2 + 5)$$

Находим градиенты функций Ф, g1, g2, g3:

$$\nabla \Phi(X) = \begin{pmatrix} 2(x_1-4) \\ 2(x_2-4) \end{pmatrix}$$

$$\nabla g_1(X) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}; \nabla g_2(X) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}; \nabla g_3(X) = \begin{pmatrix} -1 \\ -1 \end{pmatrix};$$

Записываем необходимое условие минимума функции Лагранжа:

$$\begin{pmatrix} 2(x_1-4) \\ 2(x_2-4) \end{pmatrix} + \lambda_1 \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \lambda_2 \begin{pmatrix} 0 \\ 1 \end{pmatrix} + \lambda_3 \begin{pmatrix} -1 \\ -1 \end{pmatrix} = 0$$

или

$$\begin{cases} 2(x_1-4) + \lambda_1 - \lambda_3 = 0 \\ 2(x_2-4) + \lambda_2 - \lambda_3 = 0 \end{cases}$$

Рассмотрим разные комбинации активных ограничений, для каждой комбинации найдем стационарную точку L. Координаты найденных точек записываем в таблицу

Точка						
Координаты						
Ф(Х)						

а) Положим, что ни одно из ограничений не является активным ограничением (точка лежит внутри области D). В этом случае можно положить $\lambda_1 = \lambda_2 = \lambda_3 = 0$ (напомним, что теорема Куна-Таккера не запрещает этого).

Тогда имеем стационарную точку функции А(4,4). Точка лежит вне области D и из рассмотрения исключается.

б) Пусть активным является ограничение g1, т.е. пусть x1=0.

Тогда можно положить $\lambda_2 = \lambda_3 = 0$, и $\lambda_1 = 8$ и $x_2 = 4$

Таким образом, имеем стационарную точку функции А1(0,4) в) Пусть активным является ограничение g2, т.е. пусть x2=0.

Тогда можно положить $\lambda_1 = \lambda_3 = 0$, и $\lambda_2 = 8$ и $x_1 = 4$

Таким образом, имеем стационарную точку функции A2(4,0)

г) Пусть активным является ограничение g3, т.е. пусть -x1-x2+5=0.

Тогда можно положить $\lambda_1 = \lambda_2 = 0$, и

$$\begin{cases} 2(x_1 - 4) - \lambda_3 = 0 \\ 2(x_2 - 4) - \lambda_3 = 0 \end{cases}$$

Откуда $x_1 = x_2 = 2.5$ и $\lambda_3 = -3$

Таким образом, имеем стационарную точку функции A3(2.5,2.5)

д) Пусть активными являются ограничения g1, g2, т.е. пусть x1=x2=0

Тогда можно положить $\lambda_3 = 0$

Тогда λ1=8, λ2=8.

Таким образом, имеем стационарную точку функции A4(0,0)

е) Аналогично: g1 и g3 активны

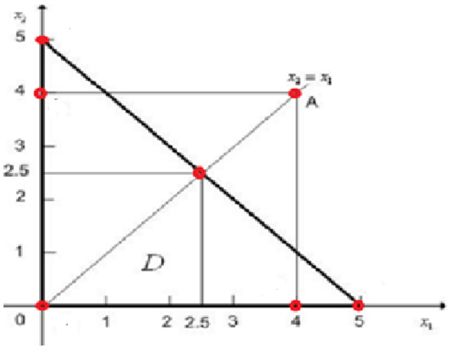
x1=0

-x1 –x2 +5=0

$\lambda_2 = 0$

X2=5 - получаем стационарную точку A5(0,5)

ж) Аналогично: g2 и g3 активны - получаем стационарную точку A6(5,0).



Легко видеть, что точки, в которых нарушаются условия регулярности ограничивающих функций, отсутствуют.

Вычисляем значения функции Ф(X) в отобранных точках (см. табл.)

Точка	A1	A2	A3	A4	A5	A6
Координаты	(0,4)	(4,0)	(2.5,2.5)	(0,0)	(0,5)	(5,0)
Ф(X)	16	16	4,5	32	17	17

Ответ: *minФ=4.5 в точке (2.5,2.5).*