Оптимизация нелинейных систем

Общая задача оптимизации заключается в отыскании экстремума целевой функции

$$Z(x_1, x_2, \dots x_n) \to \mathrm{extr}$$

n переменных, при m ограничениях, заданных в форме равенств и (или) неравенств

$$f_1(x_1, x_2, \dots x_n) = b_1,$$
$$f_2(x_1, x_2, \dots x_n) \geq b_2,$$
$$\dots\dots\dots\dots\dots\dots..$$
$$f_m(x_1, x_2, \dots x_n) \leq b_m,$$

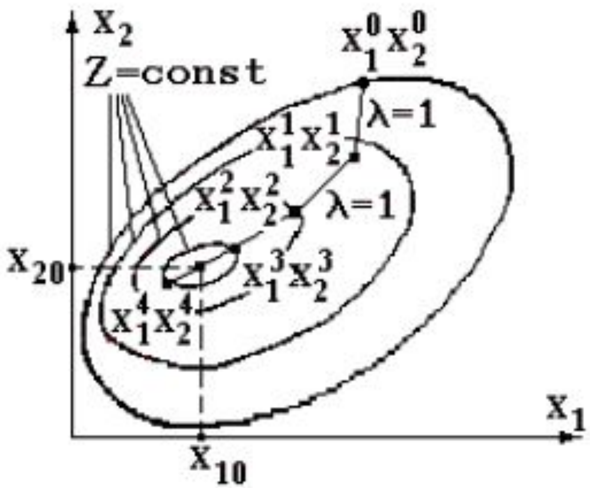
и граничных условиях, задающих диапазон изменения переменных

$$d_i \leq x_i \leq D_i, \quad i = 0, 1, 2, \dots n.$$

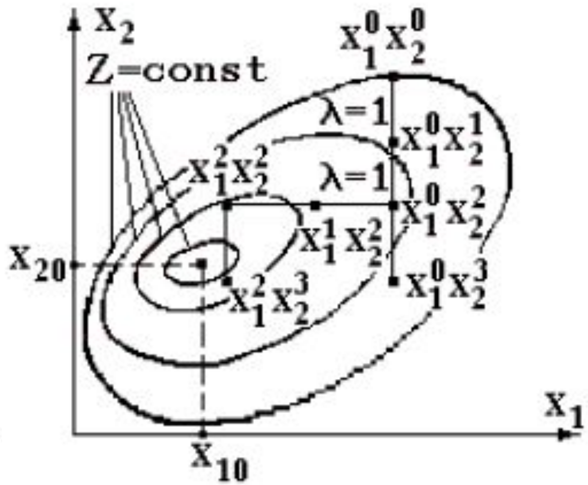
Если в математической модели оптимизационной задачи имеются нелинейные зависимости, для решения этой задачи используются методы нелинейного программирования.

Градиентные методы

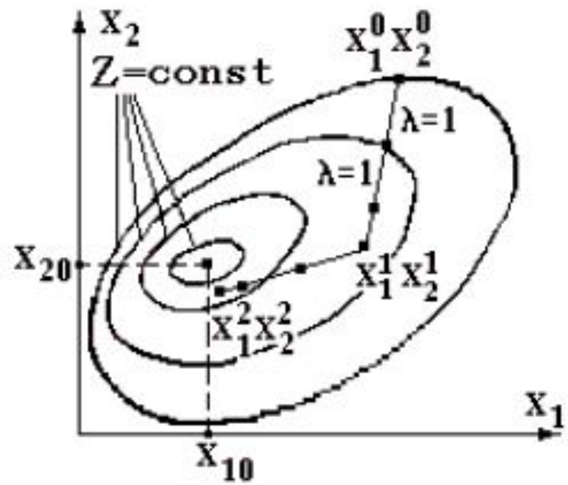
Метод с постоянным шагом



Метод покоординатного спуска



Метод скорейшего спуска



Градиентные методы

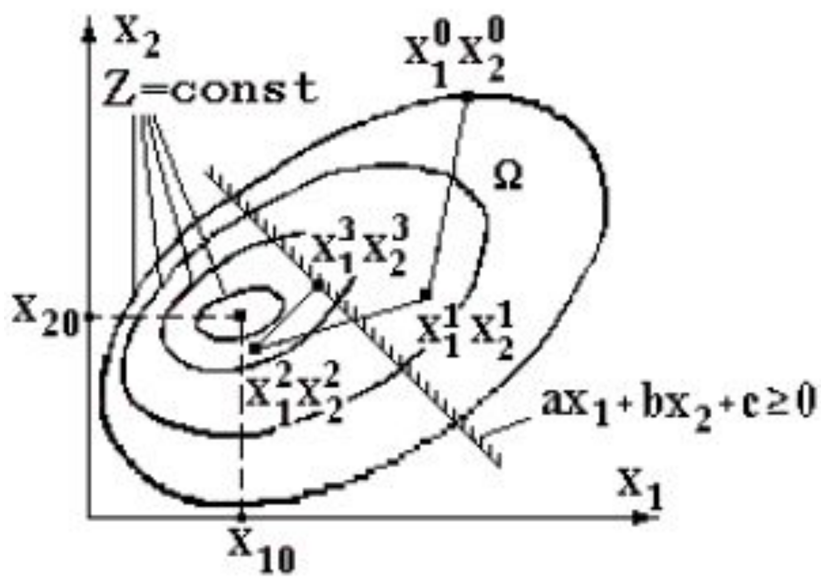
<table>
<tr><th>Метод с постоянным шагом</th><th>Метод покоординатного спуска</th><th>Метод скорейшего спуска</th></tr>
<tr><td colspan="3">1. Исходное (нулевое) приближение x_1^0, x_2^0.
2. $Z^0 = Z(x_1^0, x_2^0)$.
3. grad $Z(x_1^0, x_2^0)$</td></tr>
<tr><td>4. Шаг длиной λ в направлении −grad $Z(x_1^0, x_2^0)$ → точка (x_1^1, x_2^1)
…</td><td>4. Определение большей по модулю частной производной $\partial Z/\partial x_i$ → изменение x_i на λ до тех пор, пока Z не начнет увеличиваться → изменение другой координаты → точка (x_1^1, x_2^1)
…</td><td>4. Определение оптимальной длины шага $\lambda_{опт}$, например, параболической аппроксимацией → шаг длиной $\lambda_{опт}$ в направлении −grad $Z(x_1^0, x_2^0)$ → точка (x_1^1, x_2^1)
…</td></tr>
</table>

В результате вычислительного процесса последовательно осуществляется «спуск» к минимуму функции Z. Вычислительная процедура заканчивается, когда относительное изменение целевой функции на предыдущем i-м и последующем $(i+1)$-м шагах оказывается меньше заданной точности вычислений ε:

$$(Z_i - Z_{i+1})/Z_i \le \varepsilon.$$

Метод множителей Лагранжа

Естественно, что решение задач условной оптимизации значительно сложнее решения задач безусловной оптимизации. Естественно стремление сведения задачи условной оптимизации (поиска относительного экстремума) к более простой задаче безусловной оптимизации (поиска абсолютного экстремума). Такая процедура осуществляется в методе Лагранжа. Рассмотрим сущность этого метода.



Необходимо найти условный экстремум нелинейной функции

$$Z(x_1, x_2, \dots x_n) \to \text{extr} \tag{4.14}$$

n переменных, при m ограничениях

$$\begin{aligned} &f_1(x_1, x_2, \dots x_n) \geq b_1, \\ &f_2(x_1, x_2, \dots x_n) = b_2, \\ &\dots\dots\dots\dots\dots\dots \\ &f_m(x_1, x_2, \dots x_n) \leq b_m. \end{aligned} \tag{4.15a}$$

Ограничения-неравенства преобразуются в равенства, а свободные члены переносятся в левые части ограничений, т.е. система (4.15a) приводится к виду

$$\begin{aligned} &f_1(x_1, x_2, \dots x_n, b_1) = 0, \\ &f_2(x_1, x_2, \dots x_n, b_2) = 0, \\ &\dots\dots\dots\dots\dots\dots \\ &f_m(x_1, x_2, \dots x_n, b_m) = 0. \end{aligned} \tag{4.15}$$

В соответствии с методом Лагранжа вместо относительного экстремума функции (4.14) при ограничениях (4.15) ищется абсолютный экстремум функции Лагранжа, которая имеет следующий вид:

$$L = Z(x_1, x_2, \ldots x_n) + \lambda_1 f_1(x_1, x_2, \ldots x_n, b_1) + \lambda_2 f_2(x_1, x_2, \ldots$$

$$\ldots x_n b_2) + \ldots + \lambda_m f_m(x_1, x_2, \ldots x_n, b_m) \to \text{extr}, \qquad (4.16)$$

где $\lambda_1, \lambda_2, \ldots \lambda_m$ - неопределенные множители Лагранжа, являющиеся, как и переменные $x_1, x_2, \ldots x_n$, искомыми переменными.

Видно, что в функцию Лагранжа входит целевая функция плюс каждое ограничение, умноженное на множитель Лагранжа.

Доказано, что *относительный экстремум целевой функции (4.14) при ограничениях* (4.15) *совпадает с абсолютным экстремумом функции Лагранжа* (4.16).

Поиск абсолютного экстремумам функции (4.16) выполняется известными методами. В частности, определяются и приравниваются к нулю частные производные функции Лагранжа:

$$\begin{aligned}
&\partial L/\partial x_1 = \partial Z/\partial x_1 + \lambda_1 \partial f_1/\partial x_1 + \lambda_2 \partial f_2/\partial x_1 + \ldots + \lambda_m \partial f_m/\partial x_1 = 0,\\
&\partial L/\partial x_2 = \partial Z/\partial x_2 + \lambda_1 \partial f_1/\partial x_2 + \lambda_2 \partial f_2/\partial x_2 + \ldots + \lambda_m \partial f_m/\partial x_2 = 0,\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
&\partial L/\partial x_n = \partial Z/\partial x_n + \lambda_1 \partial f_1/\partial x_n + \lambda_2 \partial f_2/\partial x_n + \ldots + \lambda_m \partial f_m/\partial x_n = 0,\\
&\partial L/\partial \lambda_1 = f_1(x_1, x_2, \ldots x_n, b_1) = 0,\\
&\partial L/\partial \lambda_2 = f_2(x_1, x_2, \ldots x_n, b_2) = 0,\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
&\partial L/\partial \lambda_m = f_m(x_1, x_2, \ldots x_n, b_m) = 0.
\end{aligned} \qquad (4.17)$$

Последние m уравнений представляют собой ограничения (4.15) оптимизационной задачи.

Система (4.17) содержит $(m+n)$ уравнений и такое же количество неизвестных.

Решение системы (4.17) даст координаты абсолютного минимума функции Лагранжа (4.16) или относительного минимума целевой функции (4.14) при ограничениях (4.15).

Решение системы (4.17) выполняется известными методами вычислительной математики. Если система (4.17) линейная, используется, как правило, метод Гаусса. Если система (4.17) нелинейная - метод Ньютона.

Решение задач нелинейного программирования в EXCEL

***Пример** 7*. В существующей схеме электроснабжения (рис. 4.10) требуется определить мощности компенсирующих устройств $Q_{к1}$ и $Q_{к2}$ в узлах 1 и 2 исходя из условия минимума суммарных затрат на установку этих устройств и покрытие потерь активной мощности в схеме.

Исходные данные:

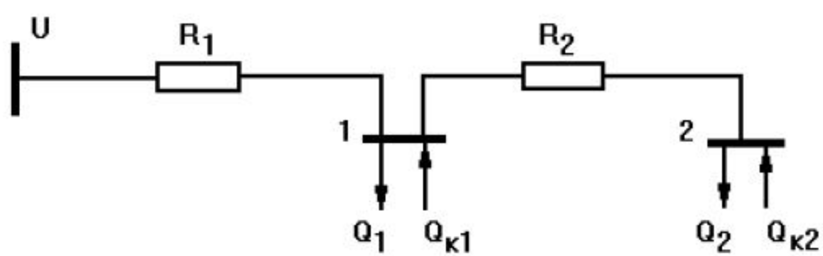
напряжение схемы U= 10 кВ;

сопротивления линий R_1=6 Ом, R_2=4 Ом;

реактивные нагрузки узлов 1 и 2 Q_1=600 квар и Q_2=800 квар;

удельные затраты на установку компенсирующих устройств z_o=0,5 у.е./квар;

удельные затраты на покрытие потерь активной мощности c_o=10 у.е./кВт.

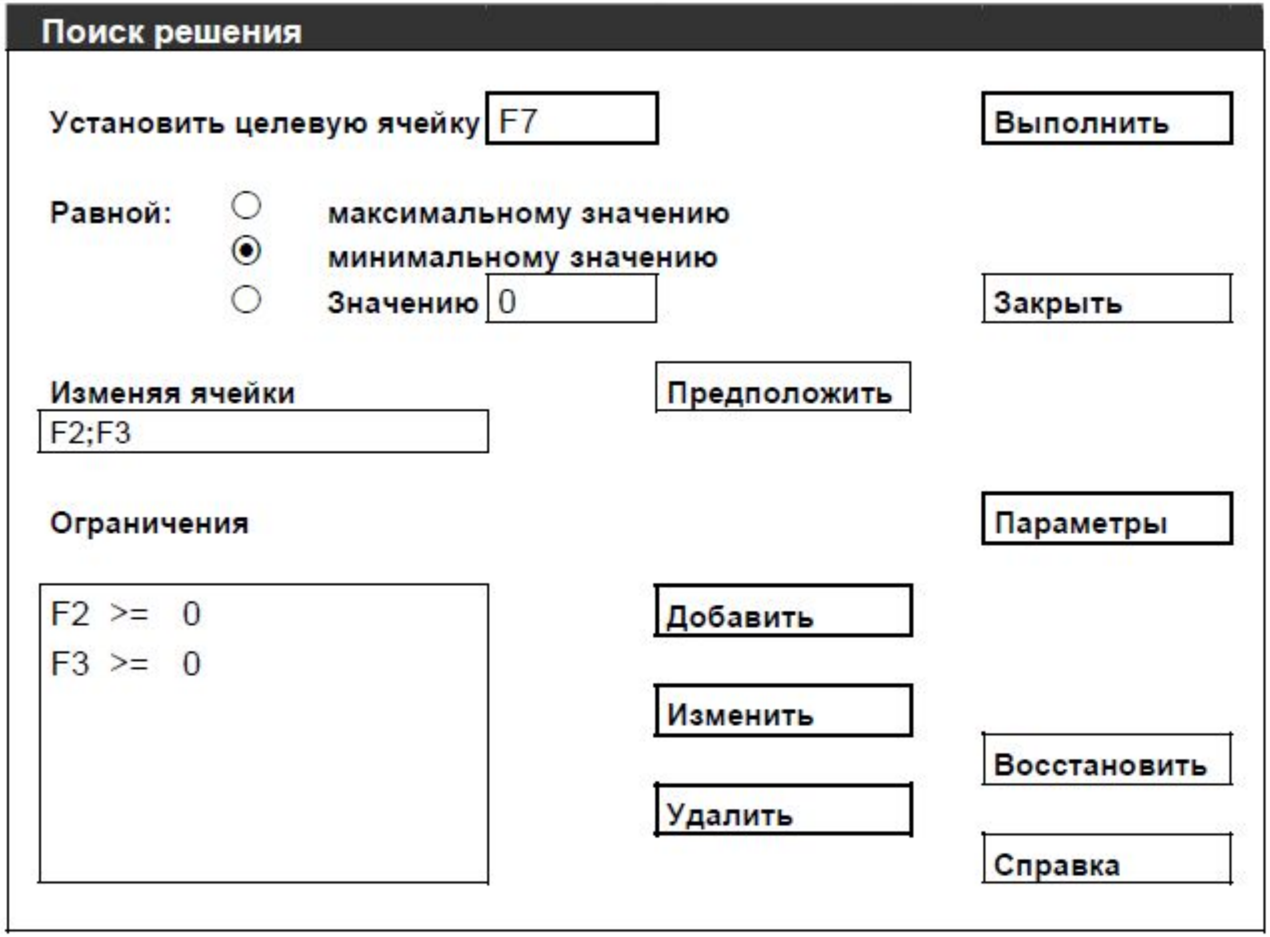


Целевая функция задачи имеет вид

$$Z=z_0(Q_{k1}+Q_{k2})+a_1(Q_1+Q_2-Q_{k1}-Q_{k2})^2+a_2(Q_2-Q_{k2})^2,$$

где $a_1=R_1\,c_o\,10^{-3}/U^2=0{,}0006$; $a_2=R_2\,c_o\,10^{-3}/U^2=0{,}0004$.

	A	B	C	D	E	F
1	**Исходные данные:**				**Переменные**	
2	Q1=	600			Qk1=	0
3	Q2=	800			Qk2=	0
4	R1=	6				
5	R2=	4				
6	U=	10			**Целевая функция**	
7	z0=	0,5			Z=	1432
8	c0=	10				
9	$a1=R1\cdot c0\cdot 10^{-3}/U^2=$	0,0006				
10	$a2=R2\cdot c0\cdot 10^{-3}/U^2=$	0,0004				
11						
12	**Ограничения:**					
13	Qk1>=0					
14	Qk2>=0					
15						



В диалоговом окне «Параметры поиска решения» следует снять флажок «**v**» с позиции «линейная модель», поскольку решается нелинейная задача.

	A	B	C	D	E	F
1	**Исходные данные:**				**Переменные**	
2	Q1=	600			Qk1=	183,3343
3	Q2=	800			Qk2=	799,9993
4	R1=	6				
5	R2=	4				
6	U=	10			**Целевая функция**	
7	z0=	0,5			Z=	595,8333
8	c0=	10				
9	$a1=R1\cdot c0\cdot 10^{-3}/U^2=$	0,0006				
10	$a2=R2\cdot c0\cdot 10^{-3}/U^2=$	0,0004				
11						
12	**Ограничения:**					
13	Qk1>=0					
14	Qk2>=0					
15						

Спасибо за внимание!