Тема 8 Нелинейное программирование

Обобщенный метод множителей Лагранжа

Пусть дана задача, в которой требуется максимизировать z=f(X)
при ограничениях
$g_i(X)\leq 0$, i=l, 2, . . ., *m*.

Ограничения X≥0, если таковые имеются, предполагаются включенными в состав данных m ограничений.

Основная идея обобщенного метода множителей Лагранжа опирается на следующий факт: если точка *безусловного,* оптимума f(X) не удовлетворяет всем ограничениям задачи, то условный оптимум должен достигаться в граничной точке области допустимых решений. Это означает, что одно или несколько ограничений из общего числа m ограничений должны выполняться как точные равенства. Соответствующая вычислительная процедура включает следующие шаги.

Шаг 1. Решить задачу без учета ограничений, т.е. максимизировать z=f(X).

Если точка полученного оптимума удовлетворяет всем ограничениям, прекратить вычисления, так как все ограничения оказываются избыточными. В противном случае положить k=1 и перейти к шагу 2.

Шаг 2. Сделать любые k ограничений активными (т. е. превратить их в равенства) и найти оптимум f(X) при наличии k активных ограничений с помощью метода множителей Лагранжа. Если полученное решение окажется допустимым по отношению к остальным ограничениям, то прекратить вычисления; *локальный* оптимум будет найден. В противном случае сделать активными другие k ограничений и повторить данный шаг. Если все подмножества, состоящие из k активных ограничений, не приводят к получению допустимого решения, перейти к шагу 3.

Шаг 3. Если k=m, прекратить вычисления; допустимых решений не существует. В противном случае положить k=k+1 и перейти к шагу 2.

С описанной выше процедурой связано важное обстоятельство, которое часто не принимается во внимание. Процедура *не* гарантирует получения глобального оптимума даже в тех случаях, когда задача имеет *единственное* оптимальное решение. Не менее важно отметить ошибочность интуитивного представления о том, что при $p<.q$ точке оптимума f(X), удовлетворяющей ***p*** ограничениям в виде равенств, всегда соответствует «лучшее» значение целевой функции по сравнению со значением, полученным при оптимизации с учетом ***q*** ограничений-равенств. Такая ситуация, имеет место только тогда, когда ***p*** ограничений образуют подмножество набора из ***q*** ограничений. *Локальный* оптимум определяется как один из всех оптимумов, полученных в результате решения оптимизационных задач с целевой функцией f(X), возникающих при учете всех комбинаций k ограничений-равенств, k=l, 2, . . ., m. Обобщенный метод множителей Лагранжа позволяет идентифицировать лишь точку локального максимума.

Следующий пример служит иллюстрацией отмеченных выше фактов.

Максимизировать $z=-(2x_1-5)^2-(2x_2-l)^{\,2}$

при ограничениях $x_1+2x_2\leq 2$, $x_1,x_2\geq 0$/

Точка безусловного оптимума находится из уравнений:

$\partial z/\partial x_1=-4(2x_1-5)=0$;

$\partial z/\partial x_2=-4(2x_2-1)=0$;

откуда $(x_1,x_2)=(5/2,1/2)$.

Поскольку найденная точка не удовлетворяет условию $x_1+2x_2\leq 2$, следует поочередно сделать одно из ограничений активным. Пусть $x_1=0$. При этом функция Лагранжа

$$L(x_1,x_2,\lambda)= -(2x_1-5)^2-(2x_2-1)^2-\lambda x_1$$

Таким образом,

$$\partial L/\partial x_1=-4(2x_1-5)-\lambda=0,$$
$$\partial L/\partial x_2=-4(2x_1-1)=0,$$
$$\partial L/\partial\lambda=-x_1=0,$$

откуда $(x_1,x_2)=(0,1/2)$.

С помощью достаточных условий можно показать, что полученное решение определяет максимум. Так как эта точка удовлетворяет всем остальным ограничениям, вычислительная процедура завершается. В результате установлено, что $(x_1,x_2)=(0,1/2)$ – локальное оптимальное решение задачи. (Заметим, что, оставшиеся ограничения $x_1+2x_2\leq 2$, $x_2\geq 0$, если поочередно делать их активными, не приводят к получению допустимых решений.) Соответствующее значение целевой функции $z=-25$.

Допустимому решению $(x_1,x_2)=(2,0)$, которое определяет точку пересечения двух прямых $x_1+2x_2=2$ и $x_2=0$, соответствует значение целевой функции $z=-2$. Оно оказывается больше значения, полученного с учетом одного активного ограничения.

Описанная выше вычислительная процедура свидетельствует о том, что при реализации обобщенного метода множителей Лагранжа не следует рассчитывать на большее, чем получение приемлемого допустимого решения задачи. Это особенно проявляется в случаях, когда целевая функция не является унимодальной. Разумеется, если в задаче имеется единственный условный оптимум, то рассматриваемый метод можно использовать для его нахождения. С этой целью следует провести сравнение безусловного оптимума и условных оптимумов, полученных с учетом каждого из всех подмножеств набора активных ограничений. При этом отдельные ограничения делаются активными поочередно, затем рассматриваются пары активных ограничений и так далее до тех пор, пока все m ограничений не станут активными. Тогда наилучший из всех таких допустимых оптимумов будет глобальным оптимумом.

В нашем случае необходимо решить семь задач, чтобы верифицировать точку глобального оптимума. Большое количество вспомогательных задач ограничивает возможности использования обобщенного метода множителей Лагранжа для решения практических задач, размерность которых, как правило, сравнительно велика.

Условия Куна — Таккера

Рассмотрим следующую задачу:

$$\text{максимизировать } z=f(X)$$

при ограничениях

$$g(X)\le 0.$$

Ограничения-неравенства можно преобразовать к виду равенств путем введения соответствующих *неотрицательных* дополнительных переменных. Таким образом, автоматически учитывая требование не отрицательности, определим $S^2_j\ge 0$ как дополнительную переменную, которую следует прибавить к левой части *i-го* ограничения $g_i(X)\le 0$. Пусть

$$S=(S_1, S_2,\dots, S_m)^T \text{ и } S^2=(S^2{}_1,S^2{}_2,\dots,S^2{}_m)^T,$$

где *m* — общее количество ограничений-неравенств. При этом функция Лагранжа записывается в следующем виде:

$$L(X, S, \lambda)=f(X)-\lambda[g(X)+S^2],$$

при заданных ограничениях $g(X)\le 0$.

Необходимым условием оптимальности является не отрицательность (не положительность) λ в задаче максимизации (минимизации).

Этот результат устанавливается следующим образом. Рассмотрим задачу максимизации. Множители λ выражают скорость изменения f по отношению к изменениям g, т.е. $\lambda=\partial f/\partial g$. Как только правая часть ограничения $g\le 0$ увеличивается и становится больше нуля, область допустимых решений расширяется. Следовательно, оптимальное значение целевой функции не может уменьшиться. Это означает, что $\lambda\ge 0$. Аналогично при увеличении правой части ограничения в задаче минимизации оптимальное значение *f* не может увеличиться, откуда следует, что $\lambda\le 0$. Если же ограничения заданы в виде равенств $g(X)=0$, то на знак λ никаких условий не накладывается.

Вычислив частные производные функции L по X, S и λ и приравняв их к нулю, получим

$$\partial L/\partial X=-\nabla f(X)-\lambda\nabla g(X)=0,$$
$$\partial L/\partial S_i=-2\lambda_i S_i=0,\ I=1,2,\ \dots,m,$$
$$\partial L/\partial\lambda=-(g(X)+S^2)=0.$$

Исследование второй группы уравнении приводит к следующим выводам.

1. Если $\lambda_i>0$ больше нуля, то $S^2{}_i=0$. Это означает, что соответствующий рассматриваемому ограничению, ресурс является дефицитным и, следовательно, исчерпан полностью (ограничение становится равенством).

2. Если $S^2_i>0$, то $\lambda_i=0$. Это означает, что i-й ресурс не является дефицитным, и, следовательно, изменение его количества не оказывает влияния на значение f($\lambda=\partial f/\partial g=0$).

Из второй и третьей групп уравнении следует, что

$\lambda_i g_i(X)=0$, i=1, 2, …,m.

Полученные условия в основном подтверждают сделанный ранее вывод, а именно если $\lambda_i>0$, то $g_i(X)=0$ или $S^2_i=0$. Аналогично, если $g_i(X)<0$, т. е. $S^2_i>0$, то $\lambda_i=0$.

Итак, **необходимые условия Куна -Таккера**, которым должны удовлетворять X и λ, определяющие стационарную точку в задаче максимизации:

$\lambda\geq 0$,

$\nabla f(X)-\lambda\nabla g(X)=0$,

$\lambda_i g_i(X)=0$, i=1, 2, …,m,

$g(X)\leq 0$.

Эти условия применимы также к задаче минимизации, если вместо условия неотрицательности λ ввести требование неположительности, необходимость которого была установлена ранее. Если же ограничения заданы в виде равенств, то на знак соответствующих множителей Лагранжа никаких условий не накладывается как в случае максимизации, так и в случае минимизации.

Алгоритмы решения нелинейных задач с ограничениями

Общую задачу нелинейного программирования можно записать в следующем виде:

максимизировать (или минимизировать) z=f(X)

при ограничениях

$g(X)\leq 0$.

Условия неотрицательности переменных $X\geq 0$ предполагаются включенными в состав заданных ограничений. Предполагается, что по крайней мере одна из функции f(X) и g(X) является нелинейной. Кроме того, потребуем, чтобы эти функции были непрерывно дифференцируемыми.

Универсальных алгоритмов решения нелинейных задач не существует, и это в первую очередь связано с исключительным «разнообразием» в поведении нелинейных функций. Наиболее общий результат, имеющий отношение к задачам нелинейного программирования, представлен условиями Куна – Таккера.

Существует ряд методов, которые можно классифицировать как прямые и непрямые. Непрямые методы, по существу позволяют получать решение нелинейных задач путем одной или нескольких линейных задач, порожденных исходной задачей. Прямые методы дают возможность построить последовательность точек, сходящуюся к точке оптимума. При этом осуществляется преобразование задачи с ограничениями в безусловной оптимизации, допускающие применение различных модификаций градиентных методов.

К непрямым методам относятся такие алгоритмы поиска опорного решения как сепарабельное, квадратичное, геометрическое и стохастическое программирование. Метод линейных комбинаций и метод последовательной безусловной максимизации, следует отнести к числу прямых методов.

Квадратичное программирование

Модель квадратичного программирования определяется следующим образом:

максимизировать (или минимизировать) $z=CX+X^{T}DX$

при ограничениях

$AX \le B$, $X \ge 0$,

где

$X=(x_1,x_2,\dots, x_n)^T$,

$C=(c_1,c_2,\dots, c_n)$,

$B=(b_1,b_2,\dots, b_m)^T$,

$$A=\begin{pmatrix} a_{11} & . & . & . & a_{1n} \\ . & & & & . \\ . & & & & . \\ . & & & & . \\ a_{m1} & . & . & . & a_{mn} \end{pmatrix}$$

$$D=\begin{pmatrix} d_{11} & . & . & . & d_{1n} \\ . & & & & . \\ . & & & & . \\ . & & & & . \\ d_{n1} & . & . & . & d_{nn} \end{pmatrix}$$

Функция X^TDX, где D–симметрическая матрица, является квадратичной формой.

Решение сформулированной задачи можно найти путем непосредственного использования необходимых условии Куна – Таккера.

В общей постановке задача записывается следующим образом:

максимизировать $z=CX+X^{T}DX$

при ограничениях

$$G(X)=\begin{pmatrix} A \\ -I \end{pmatrix}X-\begin{pmatrix} B \\ 0 \end{pmatrix} \le 0.$$

Введем множители Лагранжа

$\lambda=(\lambda_1,\lambda_2,\dots,\lambda_n)^T$, и $\mu=(\mu_1,\mu_2,\dots,\mu_n)^T$

соответствующие ограничениям $AX-B \le 0$ и $-X \le 0$ соответственно. Воспользуемся условиями Куна – Таккера:

$\lambda \ge 0$, $\mu \ge 0$,

$\nabla z-(\lambda^T, \mu^T)\nabla G(X)=0$,

$$\lambda_i (b_i-\sum_{j=1}^{n} a_{ji}x_j)=0, i=1,2,\dots, m,$$

$\mu_j x_j=0$, $j=1,2,\ldots,n$,

$AX \le B$, $-X \le 0$.

Далее получаем

$\nabla z = C + 2X^T D$,

$$\nabla G(X) = \begin{pmatrix} A \\ -I \end{pmatrix}.$$

Обозначим через $S=B-AX \ge 0$ вектор дополнительных переменных. Записанные выше условия принимают следующий вид:

$-2X^T D + \lambda^T A - \mu^T = C$,

$AX + S = B$,

$\mu_j x_j = 0 = \lambda_i S_i$, для всех i и j,

$\lambda, \mu, X, S \ge 0$.

Так как $D^T = D$, в результате перехода к транспонированным матрицам в первой группе уравнений получим

$-2XD + \lambda A^T - \mu = C^T$.

В окончательной форме запись необходимых условий выглядит так:

$$\left(\begin{array}{c|c|c|c} -2D & A^T & -I & 0 \\ \hline A & 0 & 0 & I \end{array}\right) \begin{pmatrix} X \\ \lambda \\ \mu \\ S \end{pmatrix} = \begin{pmatrix} C^T \\ B \end{pmatrix},$$

$\mu_j x_j = 0$,

$\lambda_i S_i = 0$,

$\lambda, \mu, X, S \ge 0$, для всех i и j.

Все полученные уравнения, за исключением $\mu_j x_j = 0 = \lambda_i S_i$, содержат линейные функции переменных X, λ, μ и S. Таким образом, исходная задача эквивалентна задаче нахождения решения системы линейных уравнений, удовлетворяющего дополнительным условиям $\mu_j x_j = 0 = \lambda_i S_i$. Допустимое решение, которое удовлетворяет всем этим условиям, оказывается единственным и оптимальным.

Выполнение условия $\mu_i x_j = 0 = \lambda_i S_i$ означает, что переменную S_i нельзя сделать базисной, если λ_i в базисном решении принимает положительное значение. Аналогично μ_j и x_j не могут принимать положительные значения одновременно.

Для решения полученной системы уравнений удобно применить метод искусственных переменных. Введём вспомогательную функцию

$\omega = \sum r_i$, $r_i \ge 0$, $i=1,\ldots,n$, и найдём её min.

Так как min =0 достигается при нулевых значениях r_i, то полученное опорное решение будет базисным, единственным и оптимальном, а задача в матричной форме примет вид:

$$\max z = \begin{bmatrix} 4 & 6 \end{bmatrix} \times \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} x_1 & x_2 \end{bmatrix} \times \begin{bmatrix} -2 & -1 \\ -1 & -2 \end{bmatrix} \times \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$при \quad \begin{bmatrix} 1 & 2 \end{bmatrix} \times \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \le 2, \quad X \ge 0$$

$$\left(\begin{array}{c|c|c|c|c} -2D_{n\times n} & A^T{}_{n\times m} & -I_n & 0 & I_n \\ \hline A_{m\times n} & 0 & 0 & I_m & 0 \\ \hline \multicolumn{5}{c}{\omega} \end{array} \right) \begin{pmatrix} X \\ \lambda \\ \mu \\ S \\ R \end{pmatrix} = \begin{pmatrix} C^T{}_n \\ B_m \\ 0 \end{pmatrix}$$

Так как ω записана как –1 в матрице справа, то разрешая её переменные r_i через свободные переменные получаем обычную задачу линейного программирования с необходимостью отслеживания дополнительных условий Куна – Таккера $\mu_j x_j=0=\lambda_i S_i$.

Остаётся открытым вопрос нахождения матричной формы в случае задания z полиномом. Квадратичная функция может быть представлена через её частные производные так:

$$z = \left[\frac{\partial z}{\partial x_i} \right] \times X + X^T \times \frac{1}{2} \times \left[\frac{\partial z}{\partial x_i \partial x_j} \right] \times X$$

или в операторных обозначениях $z=\nabla zX+X^T 1/2\nabla^2 zX$, откуда следует, что $C=\nabla z$, $D=1/2\nabla^2 z$.

Проверку допустимости полученного решения нетрудно осуществить путем его подстановки в систему неравенств $AX\le B$, $X\ge 0$.

Ниже приведён пример решения задачи:

Найти $\max z=4x_1+6x_2-2x_1^2-2x_1x_2-2x_2^2$

при ограничениях $x_1+2x_2\le 2$, $x_1,x_2\ge 0$.

В матричной форме задача запишется так:

Итак $C=(4,6)$, $A=(1,2)$, $B=(2)$, $S=(s_1)$, $\mu=(\mu_1, \mu_2)$ и подстановка в основное уравнение даёт

Представим матричное уравнение в виде симплекс-таблицы.

Введём искусственные переменные r_1 и r_2 в первые два уравнения, чтобы получить базис. Вспомогательную функцию $\omega= r_1+r_2$ можно непосредственно записать в строку таблицы, представив её в стандартной форме $\omega-(r_1+r_2)=0$. Заменяя базисные переменные r_1 и r_2 через их значения получим функцию ω, выраженную через свободные переменные. Таблица 1.20 готова для анализа и оптимизации.

$$\left(\begin{array}{|cc|c|cc|c|} 4 & 2 & 1 & -1 & 0 & 0 \\ 2 & 4 & 2 & 0 & -1 & 0 \\ \hline 1 & 2 & 0 & 0 & 0 & 1 \end{array} \right) \times \begin{pmatrix} x_1 \\ x_2 \\ \hline \lambda_1 \\ \hline \mu_1 \\ \mu_2 \\ \hline s_1 \end{pmatrix} = \begin{pmatrix} 4 \\ 6 \\ \hline 2 \end{pmatrix}$$

Шаг1. Строка ω указывает на выбор первого столбца, а минимальное отношение элементов столбца b к элементам столбца x_1 определяет выбор разрешающей строки r_1. Переменные x_1 и r_1 меняются местами, поскольку μ_1 свободная переменная, то выбор не нарушает дополнительных условий Куна-Таккера и допустим.

Таблица 1.20. Исходная таблица задачи квадратичного программирования.

Базисные переменные	x_1	x_2	λ_1	μ_1	μ_2	s_1	r_1	r_2	Решение b
r_1	4	2	1	-1	0	0	1	0	4
r_2	2	4	2	0	-1	0	0	1	6
s_1	1	2	0	0	0	1	0	0	2
ω	6	6	3	-1	-1	0	0	0	10

Нормируя строку r_1 и проделав равносильные преобразования с целью получения столбца единичной матрицы, получим таблицу 1.21.

Таблица 1.21. Шаг 1 задачи квадратичного программирования.

Базисные переменные	x_1	x_2	λ_1	μ_1	μ_2	s_1	r_1	r_2	Решение b
x_1	1	1/2	1/4	-1/4	0	0	1/4	0	1
r_2	0	3	3/2	1/2	-1	0	-1/2	1	4
s_1	0	3/2	-1/4	1/4	0	1	-1/4	0	1
ω	0	3	3/2	1/2	-1	0	-3/2	0	4

Шаг 2. Функция указывает на выбор второго столбца, а отношение b/ x_2 определяет строку s_1.Переменная x_2 переходит в базис, а μ_2 остаётся в числе свободных, что не нарушает дополнительных условий (таблица 1.22).

Таблица 1.22. Шаг 2 задачи квадратичного программирования.

Базисные переменные	x_1	x_2	λ_1	μ_1	μ_2	s_1	r_1	r_2	Решение b
x_1	1	0	1/3	-1/3	0	-1/3	1/3	0	2/3
r_2	0	0	2	0	-1	-2	0	1	2
x_2	0	1	-1/6	1/6	0	2/3	-1/6	0	2/3
ω	0	0	2	0	-1	-2	-1	0	2

Шаг 3. Функция ω указывает на выбор столбца λ_1 , а минимальное отношение у строки r_2 и следовательно в базис переходит λ_1, а соответствующая ей переменная s_1 остаётся в числе свободных. Операция допустима и после преобразований имеем таблицу 1.23.

Таблица 1.23. Шаг 3 задачи квадратичного программирования.

Базисные переменные	x_1	x_2	λ_1	μ_1	μ_2	S_1	r_1	r_2	Решение b

x_1	1	0	-0	-1/3	1/6	0	1/3	-1/6	1/3
λ_1	0	0	1	0	-1/2	-1	0	1/2	1
x_2	0	1	0	1/6	-1/12	1/2	-1/6	1/12	5/6
ω	0	0	0	0	0	0	-1	-1	0

Решение закончено. Вспомогательные столбцы r_1 и r_2 и строку ω можно удалить. Результат представлен в таблице 1.24.

Таблица 1.24. Оптимальное решение.

Базисные переменные	x_1	x_2	λ_1	μ_1	μ_2	S_1	Решение b
x_1	1	0	-0	-1/3	1/6	0	1/3
λ_1	0	0	1	0	-1/2	-1	1
x_2	0	1	0	1/6	-1/12	1/2	5/6

Оптимум достигнут. Дополнительные условия выполнены и решение есть $x_1=1/3$, $x_2=5/6$, $z=4\ {}^{3}/_{18}$.