9. ЗАДАЧА НА БЕЗУСЛОВНЫЕ И УСЛОВНЫЕ ЭКСТРЕМУМЫ. МЕТОД МНОЖИТЕЛЕЙ ЛАГРАНЖА

План:

1. Введение

2. Постановка задачи нелинейного программирования

3.Задачи на безусловный минимум (минимизация функции многих переменных)

4. Метод множителей Лагранжа

5. Нормальные задачи на условный минимум

6. Необходимое условие второго порядка. Достаточное условие минимума

Нелинейным программированием *называют раздел математического программирования, в котором либо целевая функция, либо ограничения, либо то и другое нелинейные.*

1. Введение

1.1. Необходимость в методах нелинейного программирования. Даже самого краткого обзора простых, но все, же поучительных моделей линейного программирования, приведенных ранее, будет достаточно для того, чтобы вызвать у читателя сомнения в действительной адекватности строго линейных моделей многим реальным ситуациям.

Легко может создаться впечатления, что при линейном подходе попросту игнорируются такие явления как: эффективность или неэффективность укрупнения операций в многопродуктовых моделях; отсутствие аддитивности объемных показателей при составлении химических смесей; влияние объема реализации на цену реализации, а следовательно, на выручку от реализации.

Отметим, что при более глубоком исследовании в ряде задач появляются связи нелинейного характера, когда с изменением одного элемента другие элементы непропорциональны первому. Например, даже простейшая транспортная задача принимает нелинейный вид, если стоимость перевозки единицы груза зависит от их общего количества.

Если у читателя и создалось такое впечатление, приведенное ниже обсуждение позволит устранить возникшие недоразумение.

Имеется много данных об очень успешном применении моделей линейного программирования в условиях нелинейности. Поскольку любая модель неизбежно оказывается лишь приближением к реальности, возникает важный вопрос: «*в каких случаях* линеаризованный вариант является адекватным отображением нелинейного явления?» Следовательно, читатель должен научиться отличать условия, в которых непосредственное применение линейного программирования приемлемо, от условий, где оно неприемлемо.

Далее следует два примера, на основе которых можно судить о вероятной неадекватности линеаризованного варианта в некоторых реальных ситуациях.

Пример 1. Рассмотрим фирму, находящуюся на начальном этапе разработки модели перспективного планирования в масштабах всей фирмы. Специалист по управлению обычно знает, что даже опытному бизнесмену трудно дать точный и детальный прогноз оптимальных объемов производства фирмы и распространения ее контроля на рынки сбыта на последующие 10 и более лет. В самом деле, применение администратором такой модели в основном обусловлено именно тем, что он понимает, как легко ошибиться при использовании только лишь интуитивных соображений в стремлении оценить влияние экономических факторов в последующие периоды. Если затраты производства и выручка от реализации зависят от объема операции нелинейно, линеаризованные догадки могут оказаться недостаточными для получения надежных ответов. (Вместе с тем при многократном применении линейная модель может оказаться применимой, если только параметры модели значительно не меняются во времени.)

Пример 2. Этот пример относится к фирме, составляющей производственную программу с помощью динамической многопродуктовой модели, отображающей существенные затраты времени на наладку станков, ограниченную мощность отдельных групп оборудования, колеблющийся спрос. Обычно в таких случаях *сущность* оптимизационной задачи состоит в варьировании различных нелинейных факторов, влияющих на принимаемые решения, относительно производственной программы. Если только специалист, разрабатывающий программу, не имеет очень хорошего представления о характере оптимального решения, любая простая линеаризация задачи, вероятно, приведет к нарушению фундаментальных принципов оптимизации.

Однако даже если в данном конкретном случае может потребоваться построение нелинейной модели, иногда удается использовать метод решения, лишь немногим отличающийся от *метода решения* для линейной модели. Наряду с этим при существенной нелинейности в связи с ее спецификой или влиянием ее на характер модели приходится применять методы оптимизации, гораздо более сложные, чем симплексный алгоритм.

1.2. Значение нелинейных моделей для управления. В настоящее время применение математического программирования в преобладающем большинстве реальных ситуаций сводится к моделям линейной аппроксимации, а не к линейным моделям в явном их виде. Однако значимость нелинейного программирования и его использования постоянно возрастает. Это обусловлено быстро растущими познаниями руководителей и специалистов в части использования математических моделей, предназначенных для подготовки решений, а также все большей доступностью компьютерных программ решения нелинейных задач большой размерности.

В большинстве случаев нелинейности, которые необходимо отобразить в моделях, относятся к одной из двух категорий:

I) эмпирически наблюдаемые соотношения, такие, как непропорциональные изменения затрат, выхода продукции, показателей качества;

II) структурно полученные соотношения, к которым относятся постулируемые экономические явления, а также выведенные математически или установленные руководством правила поведения.

Очевидно, четкое разграничение этих двух категорий невозможно, поскольку при наличии достаточных данных можно вывести структурное соотношение, лежащее в основе эмпирически наблюдаемого явления.

Пример 3. Первым примером может служить тот случай, когда на предприятии в течение ряда лет прирост выпуска продукции отстает от роста затрат труда, тогда как темпы роста количества отходов его обгоняют.

Пример 4. Вторым примером является фирма которая должна оплатить счет за электроэнергию в случае, когда расчеты ведутся по нелинейной формуле, учитывающей как среднесуточный расход, так сведения о нелинейном характере затрат из договора о ставках оплаты, заключенного с компанией, обеспечивающей энергоснабжение.

Нелинейность «встраивается» в модели программирования и в других случаях, например, в следующих:

Пример 5. ***Приготовление бензиновых смесей.*** В модели приготовления бензина определенного состава из отдельных фракций, полученных в результате перегонки нефти, обычно имеется нелинейное ограничение на октановое число смеси, поскольку эта характеристика качества нелинейно зависит от количества добавляемого к смеси тетра – этилового свинца.

Пример 6. ***Управление производственным процессом.*** В модели металлургического завода значение переменной, характеризующей температуру в доменном печи, может количеству потребного тепла и временным показателям процесса. В свою очередь каждая из этих переменных входит в другие ограничения, а также в целевую функцию.

Пример 8. ***Выручка от реализации продукции.*** Спрос на продукцию компании может существенно зависеть от цен реализации: чем ниже цена продукта, тем больше объем реализации, несмотря на аналогичное снижение цен, производимое конкурентами. Следовательно, выручка от реализации продукции не изменяется пропорционально цене, и это обстоятельство должно быть отражено в целевой функции многопродуктовой модели с помощью нелинейного слагаемого. Для иллюстрации примем, что $x(p)$ есть объем реализации, зависящий от цены p; тогда выручка от реализации равна $p \cdot x(p)$. Пусть на представляющем для нас интерес интервале изменения p функция объема реализации от цены линейна, т.е. имеет вид $x(p) = ap + b$. Тогда слагаемые в целевой функции, относящиеся к выручке от реализации, являются квадратичным относительно управляющей переменной p и имеют вид $(a\,p^2 + b\,p)$.

Пример 9. ***Размер многопродуктового заказа.*** В моделях управления запаса обычно число продуктов бывает равно одному. Однако нередко оптовый покупатель пополняет свои запасы, заказывая у одного и того же поставщика одновременно несколько видов продукции. Тем самым достигается экономия на транспортных затратах, расходах по оформлению документации и скидке на размер заказа, представляемой поставщиком. Эта ситуация может рассматриваться на основе использования модели математического программирования большой размерности, в которой затраты на пополнения запасов являются нелинейно функцией нескольких переменных – размеров заказов отдельных продуктов.

Пример 10. ***Уровень страховых запасов.*** В большинстве моделей математического программирования, используемых для общефирменного планирования, длительность отрезков планового периода редко составляет менее трех месяцев и часто превышает год и более. В таких динамических, «многопериодных» моделях обычно предусматривается условие наличия страховых запасов, которые должны выполнять компенсатора колебаний еженедельного объема реализации. В этих моделях применяется, в частности, следующий подход: уровень страхового запаса предполагается зависимым как от прогнозированного объема реализации, так и от степени использования производственных мощностей, обусловленной этим прогнозом. Так, например, пусть c - максимально возможный недельный объем производства рассматриваемого продукта, s - прогнозированный *средненедельный* объем реализации этого продукта и $n \cdot s$ - уровень страхового запаса продукта, где n - число недель, зависимое от коэффициента использования производственных мощностей s/c. Для примера предположим, что администрация приняла следующую формулу расчета n: $n = m + f(s/c)$. Тогда уровень страхового запаса представляет собой квадратичную функцию прогнозированного среднедельного уровня реализации, имеющую следующий вид: $[ms + (j/c)s^2]$. Этот уровень может входить как в ряд ограничений модели, так и в целевую функцию.

Как видно из сказанного, множество разнообразных обстоятельств приводит к нелинейной формулировке ограничений или целевых функций задач математического программирования.

2. Постановка задачи нелинейного программирования

В общем случае, **задачу нелинейного программирования** можно поставить следующим образом:

Пусть в R^n заданны функции, $f(X), g_1(X), g_2(X), \ldots, g_m(X)$ из которых хотя бы одна является нелинейной. Определим в R^n каким-нибудь образом скалярное произведение векторов X и Y, например, следующим образом:

$$Y'X = \sum_{i=1}^{n} x_i y_i = X'Y.$$

Символом $\|X\|$ будем обозначать норму вектора X, определенную $\|X\|^2 = X'X$.

Требуется найти точку, $X^0 \in R^n$ удовлетворяющую условиям:

$$g_1(X) = 0,\ g_2(X) = 0,\ \ldots,\ g_m(X) = 0,$$

такую, что

$$f(X^0) = \min_{\substack{g_i(X)=0 \\ i=\overline{1,m}}} f(X).$$

Кратко задачу нелинейного программирования в общем случае можно записать так:

$$f(X) \to \min \tag{1}$$

$$g_i(X) = 0, \quad i = 1,2,\ldots,m. \tag{2}$$

Очевидно, что задача нелинейного программирования имеет смысл лишь в том случае, когда допустимое множество K:

$$K = \{X : g_1(X) = 0,\ g_2(X) = 0, \dots, g_m(X) = 0\}$$

не пусто (система ограничений (2) совместна).

Нетрудно понять, что задача нелинейного программирования не всегда имеет решение.

Во многих случаях для доказательства существования решения задачи нелинейного программирования достаточно воспользоваться теоремой Вейерштрасса о минимуме на компакте непрерывной функции.

Задачу (1), (2) называют задачей на глобальный условный минимум в отличие от задачи на локальный условный минимум, которая состоит в следующем: найти точку X^0, удовлетворяющую ограничениям (2), такую, что при некотором достаточно малом $\varepsilon,\quad \varepsilon > 0$, для всех допустимых векторов (точек) X из ε-окрестности точки X^0, $\|X - X^0\| \le \varepsilon$, выполняется неравенство

$$f(X^0) \le f(X).$$

В других обозначениях точка X^0 - локального условного минимума – определяется равенством:

$$f(X^0) = \min_{\|X - X^0\| \le \varepsilon,\quad X \in K} f(X).$$

Замечание. Некоторые авторы используют термины абсолютный и относительный, а не глобальный и локальный условный минимум. Мы ниже будем использовать те и другие термины.

Замечание. Так как любое уравнение равносильно системе двух неравенств

$$\begin{cases} \varphi(X) \le 0, \\ -\varphi(X) \le 0, \end{cases}$$

то можно считать, что допустимое множество K задачи нелинейного программирования задается только неравенствами.

Замечание. Отметим также, что если ограничения задачи нелинейного программирования задается ограничениями типа неравенств, то их можно свести введением дополнительных переменных к ограничениям типа равенств.

Задачи нелинейного программирования образуют широкий класс оптимизационных задач. В частности, к этому классу принадлежит задача минимизации функции n переменных на всем пространстве R^n, которую называют **задачей безусловного минимума**.

Отметим, что основные результаты в нелинейном программировании получены при рассмотрении задач, в которых система ограничений линейная, а целевая функция нелинейная. Даже в таких задачах оптимальное решение может быть найдено только для узкого класса целевых функций. Рассматривают частные случаи, когда целевая функция сепарабельная (является суммой n функций $f_j(X_j)$) или квадратичная. Отметим также, что в некоторых специальных случаях эти задачи можно решить графическим методом. Мы на них будем останавливаться ниже в п. 4.1.

3. Задачи на безусловный минимум (минимизация функции многих переменных)

Приступим к исследованию функций, определенных на n-мерном пространстве R^n. Точки пространства R^n будем обозначать символом X. При операциях с вектором X будем считать его записанным в виде вектора-столбца, хотя часто для экономии места компоненты вектора будут записываться в строку. Для обозначения вектора-строки используется символ (')- транспонирование. Поэтому

$$X = \begin{pmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{pmatrix}, \qquad X' = (x_1, x_2, \dots, x_n),$$

где x_i, $i = 1,2,..,n$ - компоненты вектора X (координаты точки X).

Для дважды непрерывно дифференцируемой скалярной функции

$$f(X), \qquad X \in R^n, \text{ (т.е. } f(X) \in C^{(2)}\text{)},$$

символы $\nabla f \equiv \frac{\partial f}{\partial X}$, $\frac{\partial^2 f}{\partial X^2}$ означают

$$\nabla f \equiv \frac{\partial f}{\partial X} = \begin{pmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \\ \ldots \\ \frac{\partial f}{\partial x_n} \end{pmatrix}, \qquad \frac{\partial^2 f}{\partial X^2} = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1 \partial x_1} & \ldots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \ldots & \frac{\partial^2 f}{\partial x_2 \partial x_n} \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \ldots & \frac{\partial^2 f}{\partial x_n \partial x_n} \end{pmatrix}.$$

Если функция $g(X)$- m-мерная, то символ $\frac{\partial g}{\partial X}$ означает $n \times m$- мерную матрицу $(\nabla g_1, \nabla g_2, \ldots, \nabla g_n)$.

3.1. Постановка задачи на безусловный минимум. Пусть в R^n задана скалярная функция $f(X)$. Требуется найти точку $X^0, \quad X^0 \in R^n$, такую, что

$$f(X^0) = \min_{x \in R^n} f(X).$$

Эта задача называется **задачей на глобальный (абсолютный) минимум,** точка X^0- точкой глобального (абсолютного) минимума**. Задача на локальный (относительный) минимум** состоит в поиске точки $X^0, \quad X^0 \in R^n$, такой, что при некотором достаточно малом $\varepsilon, \quad \varepsilon > 0$, выполняется соотношение

$$f(X^0) = \min_{\|X - X^0\| \le \varepsilon, \quad X \in R^n} f(X).$$

Вопрос о существовании решения поставленных задач во многих случаях решается с помощью теоремы Вейерштрасса (п. 2).

3.2. Вспомогательные сведения из теории квадратичных форм. Пусть $A = (a_{ij})$- симметричная $n \times n$ матрица. Выражение

$$\sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i x_j = X'AX$$

называется **квадратичной формой.** Эта форма называется **знакоположительной,** если неравенство

$$X'AX \ge 0$$

выполняется для всех точек $X \in R^n$, и **определенно положительной,** если неравенство

$$X'AX > 0$$

выполняется для всех $X \in R^n$, $X \neq 0$. Аналогично определяются **знакоотрицательные** и **определенно отрицательные квадратичные формы.** С введенными квадратичными формами связаны понятия положительных и неотрицательных матриц. Симметричная матрица A называется **положительной (неотрицательной)** и обозначается $A > 0$ (соответственно, $A \geq 0$), если она служит матрицей коэффициентов определенно положительной (знакоположительной) квадратичной формы.

Рассмотрим симметрическую матрицу

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}.$$

Минор матрицы A, составленный из строк с номерами $i_1,\dots,i_p$ и столбцов $j_1,\dots,j_p$, обозначим через

$$A\begin{pmatrix} i_1,\dots,i_p \\ j_1,\dots,j_p \end{pmatrix} = \begin{vmatrix} a_{i_1 j_1} & \dots & a_{i_1 j_p} \\ \dots\dots & \dots\dots & \dots\dots \\ a_{i_p j_1} & \dots & a_{i_p j_p} \end{vmatrix}.$$

Минор называется **главным,** если $i_1 = j_1$,..., $i_p = j_p$, т. е. он составлен из строк и столбцов с одинаковыми номерами. Миноры

$$D_1 = a_{11}, \quad D_{22} = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \quad ,\dots, \quad D_n = \begin{vmatrix} a_{11} & \dots & a_{1n} \\ \dots & \dots & \dots \\ a_{n1} & \dots & a_{nn} \end{vmatrix}$$

называются **последовательными главными.**

Справедливы следующие утверждения **(критерии Сильвестра):**

1) для того чтобы матрица была положительной, необходимо и достаточно, чтобы ее последовательные главные миноры были положительны:

$$D_1 > 0 \quad ,..., \quad D_n > 0 . \tag{1}$$

2) для того чтобы матрица была неотрицательной, необходимо и достаточно, чтобы все ее главные миноры были неотрицательны:

$$A\begin{pmatrix} i_1,...,i_p \\ i_1,...,i_p \end{pmatrix} \geq 0, \quad 1 \leq i_1 \leq i_2 \leq ... \leq i_p \leq n, \quad p = 1,2,...,n . \tag{2}$$

Условий $D_1 \geq 0 \quad ,..., \quad D_n \geq 0$ недостаточно, чтобы матрица была неотрицательной. Действительно, у матрицы

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

последовательные главные миноры

$$D_1 = a_{11}, \qquad D_2 = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{11}^2 .$$

При $a_{11} = 0, \quad a_{12} = 0, \quad a_{22} < 0$, получаем $D_1 = 0, \qquad D_2 = 0$, однако в этом случае соответствующая форма $a_{11}x_1^2 + 2a_{12}x_1x_2 + a_{22}x_2^2 = a_{22}x_2^2$ не является знакоположительной (она знакоотрицательна).

Замечание. Применяя условия (1), (2) к матрице A, получаем критерии:

а) отрицательности матрицы: $(-1)^p D_p > 0, \quad p = 1,2,...,n;$

б) неположительности матрицы: $(-1)^p A\begin{pmatrix} i_1,...,i_p \\ i_1,...,i_p \end{pmatrix} \geq 0.$

Замечание. Проверку условий (2) следует начинать с построения последовательных главных миноров, ибо из неравенств (1) следует, что все главные миноры положительны.

3.3. Необходимые условия минимума. Пусть $X^0, \quad X^0 \in R^n$ - точка локального (относительного) минимума, т.е. существует такое $\varepsilon, \quad \varepsilon > 0$, что для всех $X, \quad \|X - X^0\| \leq \varepsilon$, выполняется неравенство $f(X^0) \leq f(X)$.

Теорема 1. В точке минимума X^0 гладкой функции $\left(f(X) \in C^{(1)}\right)$ выполняется условие

$$\nabla f(X^0) = 0 . \tag{3}$$

Доказательство мы опускаем, так как его можно найти в любом учебнике по математическому анализу, но по поводу гладкости (т.е. непрерывно дифференцируемости) функции хотим отметить, что каковая не излишня в данном контексте. Функция

$$f(x) = 2x^2 + x^2 \sin\left(\frac{1}{x^2}\right)$$

имеет в нуле строгий минимум и равную нулю производную, но в сколь угодно малой окрестности нуля принимает, как положительные, так и отрицательные значения (см. рис.1).

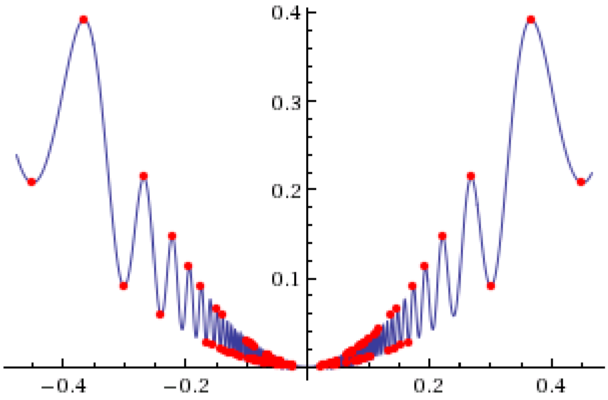

Рис.1

Условие (3) называется **необходимым условием минимума** первого порядка. Вектор $\nabla f(X^0) = \dfrac{\partial f(X^0)}{\partial X}$ - принято называть **градиентом функции** $f(X)$ в точке X^0.

В новых терминах теорема 1 утверждает: в точке локального (относительного) минимума градиент функции равен нулю.

Теорема 1 сводит поиск относительного минимума к решению уравнения

$$\nabla f(X) = 0. \tag{4}$$

Решения уравнения (4) называют стационарными точками функции $f(X)$. Смысл теоремы 1 можно выразить и таким образом: точка минимума функции является стационарной точкой функции. Обратное утверждение, конечно, неверно, ибо уравнению могут удовлетворять и точки максимума и другие точки.

Для того чтобы среди стационарных точек выделить точки минимума, нужно использовать дополнительные условия, например, необходимое условие второго порядка, которое содержится в следующем утверждении.

Теорема 2. Пусть функция $f(X)$ определена, непрерывна вместе с производными первого и второго порядков во всех точках n- мерного пространства R^n. Если X^0- точка относительного минимума, то в этой точке матрица вторых производных минимизируемой функции неотрицательна:

$$\frac{\partial^2 f(X^0)}{\partial X^2} \geq 0. \tag{5}$$

Доказательство этой теоремы также опускаем.

3.4. Достаточное условие относительного минимума. Из курса математического анализа без доказательства приведем следующее утверждение.

Теорема 3. Для того чтобы стационарная точка X^* была точкой относительного минимума дважды непрерывно дифференцируемой функций $f(X)$ ($f(X) \in C^{(2)}$), достаточно, чтобы матрица вторых производных функции $f(X)$ в точке X^* была положительной

$$\frac{\partial^2 f(X^*)}{\partial X^2} > 0. \tag{6}$$

Пример 11. Исследовать на экстремум функцию

$$f(x, y, z) = x^2 + 2xy + 4xz + 8yz + 5y^2 + 9z^2.$$

Решение

Найдем стационарные точки функции $f(x, y, z)$. Они определяются из системы уравнений

$$\begin{cases} \dfrac{\partial f}{\partial x} = 2x + 2y + 4z = 0, \\ \dfrac{\partial f}{\partial y} = 2x + 10y + 8z = 0, \\ \dfrac{\partial f}{\partial z} = 4x + 8y + 18z = 0. \end{cases}$$

Определитель этой системы

$$\begin{vmatrix} 2 & 2 & 4 \\ 2 & 10 & 8 \\ 4 & 8 & 18 \end{vmatrix} = 8 \cdot \begin{vmatrix} 1 & 1 & 2 \\ 1 & 5 & 4 \\ 2 & 4 & 9 \end{vmatrix} = 8 \cdot 16 > 0 .$$

Поэтому единственное решение однородной системы есть $x = y = z = 0$.

Итак, функция $f(x, y, z)$ имеет единственную стационарную точку $(0,0,0)$. Найдем

$$\frac{\partial^2 f(0,0,0)}{\partial X^2} = \begin{pmatrix} \dfrac{\partial^2 f(0,0,0)}{\partial x^2} & \dfrac{\partial^2 f(0,0,0)}{\partial x \partial y} & \dfrac{\partial^2 f(0,0,0)}{\partial x \partial z} \\ \dfrac{\partial^2 f(0,0,0)}{\partial y \partial x} & \dfrac{\partial^2 f(0,0,0)}{\partial y^2} & \dfrac{\partial^2 f(0,0,0)}{\partial y \partial z} \\ \dfrac{\partial^2 f(0,0,0)}{\partial z \partial x} & \dfrac{\partial^2 f(0,0,0)}{\partial z \partial y} & \dfrac{\partial^2 f(0,0,0)}{\partial z^2} \end{pmatrix} = \begin{pmatrix} 2 & 2 & 4 \\ 2 & 10 & 8 \\ 4 & 8 & 18 \end{pmatrix}.$$

Так как,

$$D_1 = 2 > 0, \quad D_2 = \begin{vmatrix} 2 & 2 \\ 2 & 10 \end{vmatrix} > 0, \quad D_3 = \begin{vmatrix} 2 & 2 & 4 \\ 2 & 10 & 8 \\ 4 & 8 & 18 \end{vmatrix} = 8 \cdot 16 > 0,$$

то, в силу критерия Сильвестра

$$\frac{\partial^2 f(0,0,0)}{\partial X^2} = \begin{pmatrix} 2 & 2 & 4 \\ 2 & 10 & 8 \\ 4 & 8 & 18 \end{pmatrix} > 0 .$$

Точка $(0,0,0)$ является точкой локального минимума функции $f(x, y, z)$, причем $f_{\min} = f(0,0,0) = 0$.

Пример 12. Исследовать на экстремум функцию

$$f(x, y, z) = 3x^3 + 6xy + y^2 + z^2 - 2z + 1.$$

Решение

Найдем стационарные точки функции $f(x,y,z)$. Они определяются из системы уравнений

$$\begin{cases} \dfrac{\partial f}{\partial x} = 9x^2 + 6y = 0, \\ \dfrac{\partial f}{\partial y} = 6x + 2y = 0, \\ \dfrac{\partial f}{\partial z} = 2z - 2 = 0. \end{cases}$$

Решив эту систему, найдем стационарные точки $(2,-6,1)$ и $(0,0,1)$. Вычислив частные производные построим матрицу

$$\frac{\partial^2 f(x,y,z)}{\partial X^2} = \begin{pmatrix} \dfrac{\partial^2 f(x,y,z)}{\partial x^2} & \dfrac{\partial^2 f(x,y,z)}{\partial x \partial y} & \dfrac{\partial^2 f(x,y,z)}{\partial x \partial z} \\ \dfrac{\partial^2 f(x,y,z)}{\partial y \partial x} & \dfrac{\partial^2 f(x,y,z)}{\partial y^2} & \dfrac{\partial^2 f(x,y,z)}{\partial y \partial z} \\ \dfrac{\partial^2 f(x,y,z)}{\partial z \partial x} & \dfrac{\partial^2 f(x,y,z)}{\partial z \partial y} & \dfrac{\partial^2 f(x,y,z)}{\partial z^2} \end{pmatrix} = \begin{pmatrix} 18x & 6 & 0 \\ 6 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$

В точке $(2,-6,1)$ имеем

$$\frac{\partial^2 f(2,-6,1)}{\partial X^2} == \begin{pmatrix} 32 & 6 & 0 \\ 6 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$

Так как,

$$D_1 = 32 > 0, \quad D_2 = \begin{vmatrix} 32 & 6 \\ 6 & 2 \end{vmatrix} = 36 > 0, \quad D_3 = \begin{vmatrix} 32 & 6 & 0 \\ 6 & 2 & 0 \\ 0 & 0 & 2 \end{vmatrix} = 72 > 0,$$

то, в силу критерия Сильвестра

$$\frac{\partial^2 f(2,-6,1)}{\partial X^2} = \begin{pmatrix} 32 & 6 & 0 \\ 6 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix} > 0.$$

Точка $(2,\,-6,\,1)$ является точкой локального минимума функции $f(x,y,z)$, причем $f_{\min} = f(2,-6,1) = -12$.

В точке $(0,0,1)$ имеем

$$\frac{\partial^2 f(0,0,1)}{\partial X^2} = \begin{pmatrix} 0 & 6 & 0 \\ 6 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$

Для исследования функции в точке $(0, 0, 1)$ нельзя использовать критерий Сильвестра, так как $D_1 = 0$. Легко видеть, что в этой точке экстремума нет. В самом деле, $f(0,0,1) = 0$, а в столь угодно малой окрестности точки $(0,0,1)$ функция принимает как положительные, так и отрицательные значения. Например, $f(0,0,\varepsilon) > 0$, если $\varepsilon > 0$ и $f(0,0,\varepsilon) < 0$, если $\varepsilon < 0$.

Если матрица $\frac{\partial^2 f}{\partial X^2}$ вырождена, ситуация принципиально усложняется. В скалярном случае о характере критической точки можно судить по первому ненулевому члену ряда Тейлора. Для функций n переменных дело обстоит совершенно иным образом по особой схеме. Эта схема позволяет свести анализ экстремали функции к исследованию стационарной точки функции меньшего числа переменных – информативной функции. Рассмотрение этой схемы исследования вырожденных стационарных точек выходит за рамки данного курса.

Предвидеть трудности (но не их масштаб) довольно легко. Понятно, что в отсутствие знакоопределенности дифференциалов порядка выше второго - поверхности их вырождения могут накладываться друг на друга весьма разнообразно, в результате чего аномалии становятся нормой.

Островок порядка характеризуется невырожденными критическими точками, в которых невырожденны соответствующие матрицы Гессе. Классическая лемма Морса гарантирует существование в некоторой окрестности критической точки - локальной системы координат $x_1, x_2, \ldots, x_n$ такой что

$$f = x_1^2 + \cdots x_l^2 - x_{l+1}^2 - \cdots - x_n^2$$

Функция такого вида называется **морсовским l-седлом.** В случае $l = 0$ 0 имеем максимум, при $l = n$ - минимум.

Лемма Морса дает по существу полную классификацию невырожденных критических точек. Для вырожденных критических точек такой классификации нет. Простой пример вырожденной критической точки **- обезьянье седло** (рис. 3), которое описывается функцией

$$f(x,y) = x^3 - 3xy^2.$$

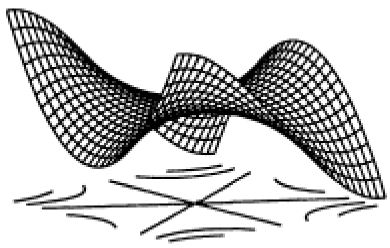

Рис. 3

Упражнения

Исследовать на экстремум функции двух переменных (1-17).

1. $f(x,y) = x^3 + 3xy^2 - 39x - 36y + 26$.

2. $f(x,y) = x^2 + xy + y^2 - 12x - 3y$.

3. $f(x,y) = 3 + 2x - y - x^2 + xy - y^2$.

4. $f(x,y) = 3x + 6y - x^2 - xy + y^2$.

5. $f = 4x^2 - 4xy + y^2 + 4x - 2y + 1$.

6. $f(x,y) = (x + y^2)e^{\frac{x}{2}}$.

7. $f(x,y) = (x^2 - 2y^2)e^{x-y}$.

8. $f(x,y) = (8x^2 - 6xy + 3y^2)e^{2x+3y}$.

9. $f(x,y) = (5 - 2x + y)e^{x^2-y}$.

10. $f(x,y) = \dfrac{x^3}{3} + 3x^2e^y - e^{-y^2}$.

11. $f(x,y) = (25 - 5x - 7y)e^{-(x^2+xy+y^2)}$.

12. $f(x,y) = 108\ln x - xy^2 + \dfrac{y^3}{3}$.

13. $f(x,y) = x^2 + y^2 - 32\ln(xy)$.

14. $f(x,y) = xy\ln(x^2 + y^2)$.

15. $f(x,y) = x + y + 4\sin x \sin y$.

16. $f(x,y) = (1 + e^y)\cos x - ye^y$.

17. Найти все стационарные точки функции $f(x,y) = x^4 + y^4 - 2x^2$ и исследовать ее на экстремум. Можно ли использовать при этом достаточное условие экстремума?

18. Может ли непрерывная дифференцируемая функция $f(x,y)$ иметь бесконечное множество максимумов и ни одного минимума?

19. Верно ли утверждение: если непрерывно дифференцируемая функция $f(x,y)$, $(x,y)\in R^2$, имеет только одну стационарную точку (x_0,y_0), в которой у нее локальный минимум, то справедливо неравенство $f(x,y)\geq f(x_0,y_0)$, $(x,y)\in R^2$?

Исследовать на экстремум функции трех переменных (21-30).

21. $f(x,y,z)=3x^3+y^2+z^2+6xy-2z+1$.

22. $f(x,y,z)=x^2+y^2+(z+1)^2-xy+x$.

23. $f(x,y,z)=8-6x+4y-2z-x^2-y^2-z^2$.

24. $f(x,y,z)=x^2+y^2-z^2-4x+6y-2z$.

25. $f(x,y,z)=x^3+y^2+z^2+6xy-4z$.

26. $f(x,y,z)=zyz(16-x-y-2z)$.

3.5.Замечание к вопросу о глобальных экстремумах. Существенный интерес нередко представляет вопрос о глобальных экстремумах. В одномерном случае локальный минимум при отсутствии других стационарных точек является одновременно глобальным минимумом. В общем случае это не так. Вот соответствующий контрпример.

Пример 13. Пусть X_0 - единственная стационарная точка дважды непрерывно дифференцируемой функции $f(X)$, $X\in R^n$, и X_0 реализует его локальный минимум. Будет, ли эта точка точкой его глобального минимума? Ответ на этот вопрос отрицателен. Соответствующий пример дает функция

$$f(x,y)=\frac{3x^2-2x^3-1}{1+y^2}+(3x^2-2x^3)e^{-y}.$$

Эта функция имеет лишь нулевую стационарную точку, которая реализует ее локальный минимум, поскольку матрица

$$\frac{\partial^2 f(0,0)}{\partial X^2}=\begin{pmatrix}12 & 0\\ 0 & 2\end{pmatrix}>0.$$

Однако, $f(0,0)=-1$, $f(2,0)=-9$.

Существуют различные признаки существования глобального минимума (максимума). Они связаны с понятиями, введения которых выходят за рамки данного курса.

Упражнение

Покажите, что если функция $f(X)$ имеет в точке $X^{\bullet}$ локальный минимум, и $\lim\limits_{\|X\|\to\infty}\|F(X)\| = \infty$, то в $X^{\bullet}$ достигается глобальный минимум $f(X)$.

4. Метод множителей Лагранжа

По условиям задачи (7), (8)составим функцию

$$F\left(X,\overline{\Lambda}\right) = \lambda_0 f(X) + \Lambda' g(X), \tag{10}$$

где $\overline{\Lambda} = (\lambda_0, \Lambda)$ есть $(m+1)$ - мерный вектор, состоящий из скаляра λ_0 и m - мерного вектора Λ.

Функцию $F\left(X,\overline{\Lambda}\right)$ называют **функцией Лагранжа** задачи (7), (8), а компоненты $\lambda_0, \lambda_1, \lambda_{2,} ..., \lambda_m$ вектора $\overline{\Lambda}$ - множителями Лагранжа.

Теорема 4 (правило множителей Лагранжа). Пусть функции

$$f(X),\ g_1(X),\ g_2(X),\ \dots,\ g_m(X)$$

определены, непрерывны и дифференцируемы на R^n. Если X^0 - точка локального условного минимума, то найдется ненулевой вектор $\overline{\Lambda}$ такой, что

$$\frac{\partial F\left(X^0,\overline{\Lambda}\right)}{\partial X} = 0.$$

Замечание. Если $\overline{\Lambda}$, удовлетворяет правилу множителей Лагранжа, то, в силу однородности по $\overline{\Lambda}$ функции Лагранжа, правилу множителей удовлетворяет и вектор -$\overline{\Lambda}$. Поэтому знак одной компоненты вектора $\overline{\Lambda}$ можно выбрать заранее. Для определенности задают знак множителя $\lambda_0 : \lambda_0 \geq 0$. В уточненной формулировке правило множителей утверждает о существовании чисел $\lambda_0 \geq 0, \lambda_1, ..., \lambda_m$, не равных одновременно нулю, таких что $\dfrac{\partial F\left(X^0,\overline{\Lambda}\right)}{\partial X} = 0$.

В приложениях часто используется функция Лагранжа,

$$F(X,\Lambda) = f(X) + \Lambda' g(X), \tag{11}$$

которую будем называть **нормальной функцией Лагранжа**. Функция (11) получается из (10) при $\lambda_0 = 1$. Нормальная функция Лагранжа проще функции Лагранжа, удобнее для вычислений. Однако правило множителей для нее не всегда верно.

Пример 14. $n = 2$, $m = 1$, $f(x) = x_1$, $g(x) = x_1^3 - x_2^2$. Допустимые точки удовлетворяют уравнению $x_1^3 - x_2^2 = 0$ и лежат на полукубической параболе.

Ясно, что точка $x_1^0 = x_2^0 = 0$ есть точка условного минимума. Составим нормальную функцию Лагранжа

$$F(X,\lambda) = f(X) + \lambda g(X) = x_1 + \lambda(x_1^3 - x_2^2).$$

Правило множителей приводит к уравнениям

$$\frac{\partial F}{\partial x_1} = 1 + 3\lambda x_1^2 = 0, \qquad \frac{\partial F}{\partial x_2} = -2\lambda x_2 = 0.$$

Точка минимума $x_1^0 = x_2^0 = 0$ не удовлетворяет этим уравнениям, т. е. правило множителей с нормальной функцией Лагранжа в данной задаче не справедливо.

Множители λ_0, Λ, при которых выполняется правило множителей в точке X^0, назовем множителями, соответствующими точке X^0. Точке X^0 может соответствовать несколько систем множителей Лагранжа.

5. Нормальные задачи на условный минимум

Рассматриваются задачи на условный минимум, которые не сводятся к перебору конечного числа допустимых точек.

5.1. Нормальные точки минимума и обыкновенные допустимые точки. Если среди систем множителей Лагранжа, соответствующих точке условного минимума X^0 задачи

$$f(X) \to \min, \quad g(X) = 0,$$

нет множителей $\lambda_0, \lambda_1, \lambda_2, ..., \lambda_m$ с $\lambda_0 = 0$, то точка называется **нормальной точкой минимума;** а сама задача - **нормальной задачей на условный минимум**.

Если X^0- нормальная точка минимума, то соответствующие ей множители можно разделить на $\lambda_0 \neq 0$ и считать, что множители Лагранжа имеют вид $\lambda_0 = 1, \lambda_1, \lambda_2, ..., \lambda_m$.

Для нормальной точки соответствующие множители $1, \lambda_1, \lambda_2, ..., \lambda_m$ существуют и определяются единственным образом.

Действительно, если допустить еще одну систему множителей $1, \tilde{\lambda}_1, \tilde{\lambda}_2, ..., \tilde{\lambda}_m$, $\left(\tilde{\lambda}_i \neq \lambda_i\right)$, то получаем

$$\frac{\partial f\left(X^0\right)}{\partial X} + \lambda_1 \frac{\partial g_1\left(X^0\right)}{\partial X} + ... + \lambda_m \frac{\partial g_m\left(X^0\right)}{\partial X} = 0,$$

$$\frac{\partial f\left(X^0\right)}{\partial X} + \tilde{\lambda}_1 \frac{\partial g_1\left(X^0\right)}{\partial X} + ... + \tilde{\lambda}_m \frac{\partial g_m\left(X^0\right)}{\partial X} = 0.$$

Вычитая одно уравнение из другого, приходим к равенству

$$\mu_1 \frac{\partial g_1\left(X^0\right)}{\partial X} + ... + \mu_m \frac{\partial g_m\left(X^0\right)}{\partial X} = 0,$$

где $\mu_i = \lambda_i - \tilde{\lambda}_i \neq 0$. Полученное равенство означает, что точке X^0 соответствует еще одна система множителей $\mu_0 = 0, \mu_1, ..., \mu_m$, и это противоречит определению нормальности точки X^0. Следовательно, система множителей Лагранжа $1, \lambda_1, \lambda_2, ..., \lambda_m$единственна. Оказывается, характер точки минимума (нормальна она или нет) определяется полностью ограничениями задачи.

Точку $\tilde{X}$ назовем **обыкновенной допустимой точкой**, если $g\left(\tilde{X}\right) = 0$ и векторы

$$\frac{\partial g_1\left(\tilde{X}\right)}{\partial X}, \frac{\partial g_2\left(\tilde{X}\right)}{\partial X}, ..., \frac{\partial g_m\left(\tilde{X}\right)}{\partial X},$$

вычисленные в этой точке, линейно независимы.

Справедлива следующая теорема.

Теорема 5. Точка X^0 является нормальной в том и только в том случае, когда она обыкновенная допустимая точка.

Доказательство. Необходимость. Пусть точка X^0 - нормальная:

$$\frac{\partial f\left(X^0\right)}{\partial X}+\lambda_1\frac{\partial g_1\left(X^0\right)}{\partial X}+\ldots+\lambda_m\frac{\partial g_m\left(X^0\right)}{\partial X}=0,$$

но не является обыкновенной, т.е. для некоторого вектора $\mathrm{M}=(\mu_1,\ldots,\mu_m)$, $\mathrm{M}\neq 0$, выполняется равенство

$$\mu_1\frac{\partial g_1\left(X^0\right)}{\partial X}+\ldots+\mu_m\frac{\partial g_m\left(X^0\right)}{\partial X}=0.$$

Отсюда следует, что точке X^0 соответствует система множителей Лагранжа $\mu_0=0,\mu_1,\ldots,\mu_m=0$, чего не может быть по определению нормальной точки X^0.

Необходимость доказана.

Достаточность. Пусть допустимая точка X^0 обыкновенная. Если предположить, что она не является нормальной, то найдутся множители $\lambda_0=0,\lambda_1,\lambda_2,\ldots,\lambda_m;$ $\Lambda\neq 0$, такие, что

$$\lambda_0\frac{\partial f\left(X^0\right)}{\partial X}+\lambda_1\frac{\partial g_1\left(X^0\right)}{\partial X}+\ldots+\lambda_m\frac{\partial g_m\left(X^0\right)}{\partial X}=0.$$

Поскольку первое слагаемое слева равно нулю, то оставшаяся часть равенства означает, что векторы $\frac{\partial g_1\left(X^0\right)}{\partial X},\frac{\partial g_2\left(X^0\right)}{\partial X},\ldots,\frac{\partial g_m\left(X^0\right)}{\partial X}$ линейно зависимы. Это противоречит определению обыкновенной точки. Теорема доказана.

5.2. Правило множителей для нормальной задачи на условный минимум. Как следует из определения нормальных задач на условный минимум, при их исследовании достаточно пользоваться нормальными функциями Лагранжа. В этом случае правило множителей верно. Более того, оно упрощается и может быть записано в симметричном виде.

Теорема 6. Если точка X^0 доставляет решение нормальной задаче на условный минимум, то существует m-мерный вектор Λ такой, что выполняются равенства

$$\frac{\partial F\left(X^0,\Lambda\right)}{\partial X}=0,\qquad \frac{\partial F\left(X^0,\Lambda\right)}{\partial \Lambda}=0, \tag{12}$$

где $F(X,\Lambda)=f(X)+\Lambda' g(X)$.

Для доказательства утверждения достаточно заметить, что первое равенство в (12) есть следствие правила множителей для рассматриваемого нормального случая, а второе эквивалентно условию $g\left(X^0\right)=0$, которое должно, очевидно, выполняться для допустимой точки X^0.

Соотношения (12) представляют собой систему из $n+m$ уравнений относительно $n+m$ неизвестных X^0,Λ. В анормальных задачах правило множителей сводит исходную задачу к решению $n+m$ уравнений относительно $n+m+1$ неизвестных X^0,λ_0,Λ.

Если сравнить метод исключения с правилом множителей, то видно, что в первом методе исходная задача на условный минимум функции n переменных сводится к задаче безусловной минимизации функции $n-m$ переменных, во втором методе исходная задача сводится к условиям стационарности функции Лагранжа от $n+m$ аргументов. Следует подчеркнуть, что точка условного минимума X^0, вообще говоря, не является точкой минимума функции Лагранжа.

Пример 15. $n=2,\quad m=1,\quad f\left(X\right)=x_2,\quad g\left(X\right)=-x_1^2+x_2$. В обыкновенной точке $x_1^0=x_2^0=0$ достигается условный минимум. Правило множителей приводит к уравнениям

$$\frac{\partial F}{\partial x_1}=-2\lambda x_1=0,\qquad \frac{\partial F}{\partial x_2}=1+\lambda=0,$$

откуда $\lambda=-1$. Функция Лагранжа $F\left(X,-1\right)=x_1^2$ в точке $x_1=0$, соответствующей точке условного минимума, также достигает минимума.

Пример 16. $n=2,\quad m=1,\quad f\left(X\right)=x_2^3,\quad g\left(X\right)=-x_1^2+x_2$. Поскольку

$$\frac{\partial g}{\partial x_{11}}=-2x_1,\qquad \frac{\partial g}{\partial x_2}=1,$$

то

$$\frac{\partial g}{\partial X}\neq 0$$

и любая допустимая точка – обыкновенная, т.е. задача нормальная и

$$F\left(X,\lambda\right)=x_2^3+\lambda\left(x_2-x_1^2\right).$$

Правило множителей приводит к уравнениям

$$-2\lambda x_1 = 0, \quad 3x_2^2 + \lambda = 0, \quad x_2 - x_1^2 = 0 .$$

Точке условного минимума $x_1^0 = x_2^0 = 0$ соответствует $\lambda = 0$. Функция Лагранжа с этим значением $F(X,0) = x_2^3$ в точке $x_2^0 = 0$ не достигает минимума. Точка $x_2^0 = 0$ является точкой перегиба функции.

Пример 17. $n = 2, \quad m = 1, \quad f(X) = -x_2^2, \quad g(X) = x_2$. Точка $x_1^0 = x_2^0 = 0$ является нормальной точкой условного минимума. Правило множителей для функции $F(X,\lambda) = -x_2^2 + \lambda x_2$ приводит к уравнениям

$$\frac{\partial F}{\partial x_1} = 0, \qquad \frac{\partial F}{\partial x_2} = -2x_2^0 + \lambda = 0 ,$$

из которых следует, что $\lambda = 0$. При этом значении множителя функция Лагранжа $F(X,0) = -x_2^2$ достигает не минимума, а максимума.

5.3. Лемма о включении. Место нормальных задач в теории минимизации функций с дополнительными ограничениями определяется тем, что нормальные задачи на условный минимум не могут быть тривиальными, т.е. множество их допустимых векторов состоит не из конечного числа точек, а из бесконечного. Это следует из метода исключения. Такой же смысл можно придать, и следующей лемме, ярко отражающей существо классических методов исследования задач на минимум.

Лемма 3. Пусть X^*- обыкновенная допустимая точка, $g(X^*) = 0$, где $g(X)$- гладкая функция. Тогда для n- мерного вектора Y, лежащего в гиперплоскости

$$\frac{\partial g'(X^*)}{\partial X} Y = 0 , \tag{13}$$

найдется n- мерная функция $h(\beta)$ такая, что в окрестности точки $\beta = 0$ выполняется тождество

$$g(h(\beta)) \equiv 0$$

и

$$h(0) = X^*, \dot{h}(0) = Y \quad \left(\dot{h}(0) = \frac{dh(0)}{d\beta} \right).$$

Замечание. Найденная функция $h(\beta)$ по теореме о неявных функциях дифференцируема столько раз, сколько раз дифференцируема функция $g(X)$.

Геометрический смысл леммы можно пояснить следующим образом. Совокупность уравнений $g_1(X) = 0, ..., g_m(X) = 0, \quad m < n$, задает $(m-n)$-мерное многообразие. В частности, при $m - n = 2$ это будет поверхность.

Уравнение

$$\frac{\partial g'(X^*)}{\partial X} Y = 0$$

задает касательную гиперплоскость к этой поверхности в точке $X = X^*$. Лемма о включении утверждает, что какой бы вектор Y ни взять из касательной плоскости, на поверхности можно провести линию, которая исходит из точки $X = X^*$ с касательной, содержащей вектор Y.

По доказанной лемме, точку минимума нормальной задачи можно погрузить в семейство допустимых точек, зависящее от параметров β, Y. Поэтому нормальная задача на условный минимум не может оказаться тривиальной. Анормальные же задачи могут быть и тривиальными.

Пример 18. $n = 4, \quad m = 1, \quad f(X) = x_1, \quad g(X) = x_1^2 + x_2^2 = 0$. Множество допустимых точек состоит лишь из одной точки $x_1 = x_2 = 0$, поэтому задача минимизации вырождается: минимизировать функцию $f(X)$ на других точках не имеет смысла.

5.4. Экономическая интерпретация множителей Лагранжа. В конкретных прикладных вопросах множители Лагранжа имеют содержательную интерпретацию. Так, в механике множители Лагранжа задают реакции связей, а в экономике – **цены на продукты производства.**

Ниже будем обсуждать экономическую интерпретацию множителей Лагранжа.

Пример 19. Пусть производственные функции каждого из двух продуктов зависят от двух (одних и тех же) факторов. Суммарное количество каждого фактора фиксировано. Пусть заданы цены продуктов. При каких условиях доход от выпуска будет максимальным?

Обозначим через x_1, x_2 объемы выпуска первого и второго продуктов, соответственно. Пусть x_3, x_4 - объемы факторов, использованные при производстве первого продукта с производственной функцией $x_1 = \varphi(x_3, x_4)$. Аналогично, $x_2 = \psi(x_5, x_6)$. Предполагается, что x_3 и x_5 представляют один и тот же фактор, так же как и x_4 и x_6. Задача состоит в следующем:

$$\begin{aligned}
&f(X) = p_1 x_1 + p_2 x_2 \to \max,\\
&x_1 - \varphi(x_3, x_4) = 0, \quad x_3 + x_5 - k_1 = 0,\\
&x_2 - \psi(x_5, x_6) = 0, \quad x_4 + x_6 - k_2 = 0.
\end{aligned}$$

Функция Лагранжа в этом случае запишется в виде

$$F(X, \Lambda) =$$
$$= p_1 x_1 + p_2 x_2 - \lambda_1(x_1 - \varphi(x_3, x_4)) - \lambda_2(x_2 - \psi(x_5, x_6)) - \lambda_3(x_3 + x_5 - k_1) - \lambda_4(x_4 + x_6 - k_2).$$

Приравнивая частные производные нулю, получим шесть уравнений:

$$p_1 = x_{1,} \quad p_2 = x_2, \ \lambda_1 \frac{\partial \varphi}{\partial x_3} = \lambda_3, \quad \lambda_1 \frac{\partial \varphi}{\partial x_4} = \lambda_4, \quad \lambda_2 \frac{\partial \psi}{\partial x_5} = \lambda_3; \quad \lambda_2 \frac{\partial \psi}{\partial x_6} = \lambda_4.$$

Отсюда легко получить условия на выпуск, максимизирующий доход:

$$p_1 \frac{\partial \varphi}{\partial x_3} = p_2 \frac{\partial \psi}{\partial x_5} = \lambda_3, \qquad p_1 \frac{\partial \varphi}{\partial x_4} = p_2 \frac{\partial \psi}{\partial x_6} = \lambda_4.$$

Следовательно, стоимость предельного продукта по каждому фактору будет одна и та же в обеих отраслях. Заметим, что все множители Лагранжа оказались равными ценам. Если в приведенном случае имеет место конкурентное ценообразование, то λ_1, λ_2 - цены продуктов, а λ_3, λ_4 - цены факторов.

Таким образом, в примере, приведенном выше, множители Лагранжа оказались равными ценам. Такими же свойствами, как мы помним, обладают двойственные переменные в теории линейного программирования.

Ниже будет предложена формальная интерпретация множителей Лагранжа.

Рассмотрим стандартную задачу условной оптимизации. Решим ее с помощью метода множителей Лагранжа, который дает оптимальные векторы X^*, Λ^*. Пусть i-е ограничение имеет вид $g_i(x) = b_i$.

Первоначально полагалось, что $b_i = 0$. Исследуем здесь влияние малого ослабления ограничения.

Обозначим через V^* оптимальное значение целевой функции ($V^* = f(X^*)$). Малое ослабление i-го ограничения приводит к малым изменениям оптимальных значений переменных. Однако предполагается, что условия оптимальности по-прежнему удовлетворяются, так что новое состояние, достигаемое в результате ослабления ограничений, также оптимально. Влияние ослабления на оптимальное значение целевой функции определяется формулой

$$\frac{\partial V^*}{\partial b_i} = \sum_j \frac{\partial f(x^*)}{\partial x_j} \frac{\partial x_j}{\partial b_i}. \tag{14}$$

Из ограничений имеем

$$\sum_j \frac{\partial g_i(x^*)}{\partial x_j} \frac{\partial x_j}{\partial b_i} = \begin{cases} 0, & k \neq i, \\ 1, & k = i. \end{cases} \tag{15}$$

Умножим k-е равенство в (15) на λ_k^* и просуммируем по k. Получим

$$\sum_k \sum_j \lambda_k^* \frac{\partial g_i(x^*)}{\partial x_j} \frac{\partial x_j}{\partial b_i} = \lambda_i^*.$$

Вычтем это выражение из (14). Получим

$$\frac{\partial V^*}{\partial b_i} = \lambda_i^* + \sum_j \left[\frac{\partial f(x^*)}{\partial x_j} - \sum_k \lambda_k \frac{\partial g_i(x^*)}{\partial x_j} \right] \frac{\partial x_j}{\partial b_i}.$$

Выражение справа в скобках в силу условия оптимальности равно нулю, поэтому

$$\frac{\partial V^*}{\partial b_i} = \lambda_i^*.$$

Таким образом, λ_i^* соответствует маргинальной (предельной) скорости изменения целевой функции относительно малого ослабления i-го ограничения при условии, что все остальные ограничения неизменны. Эта интерпретация аналогична интерпретации двойственных переменных в теории линейного программирования.

В типичных экономических приложениях ограничения могут задаваться лимитами на ресурсы, а целевая функция - некоторым индексом общественного благосостояния. Тогда оптимальные множители Лагранжа соответствовали бы маргинальным (предельным) общественным оценкам ресурсов. В примере, рассмотренном выше, множители соответствуют ценам на продукты и на факторы. Цены на факторы соответствуют предельным оценкам для фиксированного предложения факторов. Чему же соответствуют цены на продукты? Оказывается, они соответствуют маргинальным оценкам производственных функций, выступающих в качестве ограничений, или параметрам эффективности производственных функций.

Таким образом, множители Лагранжа в общей задаче оптимизации, играют роль двойственных переменных в линейном программировании, и сводится к ним, если общая задача линейна.

В экономических задачах можно, поэтому интерпретировать множители Лагранжа так же, как двойственные переменные.

6. Необходимое условие второго порядка. Достаточное условие минимума

До сих пор в задачах нелинейного программирования рассматривались необходимые условия первого порядка, которые используют лишь значения первых производных. Этой информации может оказаться недостаточно для различения точек, не являющихся точками минимума. Ниже приводятся, и доказываются более тонкие условия минимума, основанные на свойствах вторых производных.

6.1. Необходимое условие второго порядка

Теорема 7. Пусть функции $f(X)$, $g(X)$ определены, непрерывны и дважды дифференцируемы в окрестности точки $X = X^0$. Если X^0 - нормальная точка условного минимума, а Λ- соответствующий ей множитель Лагранжа, то квадратичная форма

$$Y'\frac{\partial^2 f\left(X^0,\Lambda\right)}{\partial X^2}Y$$

знакоположительна на гиперплоскости

$$\frac{\partial g'\left(X^0\right)}{\partial X}Y=0, \tag{16}$$

т.е. для всех Y, удовлетворяющих уравнению (16), выполняется неравенство

$$Y'\frac{\partial^2 f\left(X^0,\Lambda\right)}{\partial X^2}Y\geq 0.$$

Доказательство опускаем.

Без предположения о нормальности точки X^0 теорема 7 может вырождаться.

Пример 20. $n=2,\quad m=1,\quad f(X)=x_1^4+x_2^4,\quad g(X)=x_1x_2$. Допустимые точки заполняют оси координат $x_1=0,\quad x_2=0$. Точка $x_1^0=0,\quad x_2^0=0$ является точкой условного минимума. Правило множителей приводит к уравнениям

$$\frac{\partial F\left(X^0,\overline{\Lambda}\right)}{\partial x_1}=4\lambda_0(x_1^0)^3+\lambda x_2^3=0,\quad \frac{\partial F\left(X^0,\overline{\Lambda}\right)}{\partial x_2}=4\lambda_0(x_2^0)^3\lambda x_1^0=0,$$

из которых следует, что любые числа $\lambda_0,\quad \lambda$ являются множителями, соответствующими точке минимума. Матрица вторых производных функции Лагранжа имеет вид

$$\frac{\partial^2 F\left(X^0,\overline{\Lambda}\right)}{\partial X^2}=\begin{pmatrix}12\lambda_0(x_1^0)^2 & \lambda\\ \lambda & 12\lambda_0(x_2^0)^2\end{pmatrix}=\begin{pmatrix}0 & \lambda\\ \lambda & 0\end{pmatrix}.$$

Уравнению

$$\frac{\partial g'(X^0)}{\partial X}Y=0\cdot y_1+0\cdot y_2=0$$

удовлетворяют все точки плоскости. Квадратичная форма $2\lambda y_1y_2$ неотрицательна лишь при $\lambda=0$.

6.2. Достаточное условие минимума. Допустимую точку X^* назовем **условно - стационарной точкой функции** $f(X)$, если найдется m- мерный вектор Λ такой, что

$$\frac{\partial F\left(X^*,\Lambda\right)}{\partial X}=0,$$

где $F(X,\Lambda)=f(X)+\Lambda' g(X)$.

Теорема 8. Пусть функции $f(X)$, $g(X)$ определены, непрерывны и дважды дифференцируемы в окрестности точки $X=X^*$. Для того чтобы условно – стационарная точка была точкой относительного условного минимума, достаточно, чтобы квадратичная форма

$$Y'\frac{\partial^2 f\left(X^*,\Lambda\right)}{\partial X^2}Y$$

была определенно положительной на гиперплоскости

$$\frac{\partial g'\left(X^*\right)}{\partial X}Y=0. \qquad (17)$$

Замечание. Из доказательства теоремы следует, что X^*- точка строгого относительного условного минимума, т.е. найдется такое ε_2, что для всех точек $X,\quad X\neq X^*,\quad X\in B_\varepsilon\quad 0<\varepsilon<\varepsilon_2$, выполняется неравенство $f\left(X^*\right)<f(X)$.

Пример 21. Найти параметры цилиндрической цистерны, которая при заданной площади поверхности S_0 имеет объем.

Решение. Площадь поверхности цистерны равна $S=2\pi x_2^2+2\pi x_1x_2$. Объем цистерны равен $V=-\pi x_1x_2^2$. Введя функцию

$$g(x_1,x_2)=\pi x_2^2+2\pi x_1x_2-S_0,$$

исходную задачу сведем к задаче

$$f(X)\to\min,\qquad g(X)=0. \qquad (18)$$

Полученная задача не эквивалентна исходной, ибо не учтены дополнительные ограничения

$$x_1\geq 0, x_2\geq 0, \qquad (19)$$

вытекающие из физической сущности этих переменных.

Однако будем решать задачу в форме (18), в случае необходимости отбрасывая решения, не удовлетворяющие ограничениям (19). Если задача (18) имеет решение, то искомое, очевидно, будет среди локальных минимумов задачи (18). Но без учета (19) задача (18) не имеет решения, ибо при $x_1 \to \infty,\ \ x_2 \to -\infty$ имеем $f(X) \to -\infty$. Значит, доказанные необходимые условия минимума нельзя применить. И все же имеющегося материала достаточно для решения поставленной задачи. Для этого нужно найти условно-стационарные точки и к ним применить достаточное условие минимума.

Составляем нормальную функцию Лагранжа

$$F(X,\lambda) = -\pi x_1 x_2^2 + \lambda[2\pi x_2^2 + 2\pi x_1 x_2 - S_0].$$

Условно-стационарные точки удовлетворяют уравнениям:

$$\frac{\partial F}{\partial x_1} = -2\pi x_2^2 + 2\pi\lambda x_2 = 0,$$

$$\frac{\partial F}{\partial x_2} = -2\pi x_1 x_2 + 4\pi\lambda x_2 + 2\pi\lambda x_1 = 0,$$

$$2\pi x_2^2 + 2\pi x_1 x_2 - S_0 = 0.$$

Отсюда

$$x_1 = 4\lambda, \quad x_2 = 2\lambda, \quad \lambda = \pm\sqrt{\frac{S_0}{24\pi}}.$$

Выберем $\lambda > 0$. Подсчитаем матрицу

$$\frac{\partial^2 F}{\partial X^2}\begin{pmatrix} 0 & -2\pi x_2 + 2\pi\lambda \\ -2\pi x_2 + 2\pi\lambda & -2\pi x_1 + 4\pi\lambda \end{pmatrix} = \begin{pmatrix} 0 & -2\pi\lambda \\ -2\pi\lambda & -4\pi\lambda \end{pmatrix}$$

Уравнение гиперплоскости, на которой проверяется квадратичная форма:

$$\frac{\partial g}{\partial x_1} y_1 + \frac{\partial g}{\partial x_2} y_2 = 2\pi x_2 y_1 + [4\pi x_2 + 2\pi x_1] y_2 = 4\pi\lambda y_1 + 16\pi\lambda y_2 = 0. \qquad (20)$$

Квадратичная форма с матрицей $\frac{\partial^2 F}{\partial X^2}$ имеет вид $-4\pi\lambda y_1 y_2 - 4\pi\lambda y_2^2$

и на гиперплоскости (20) переходит в выражение $12\pi\lambda y_2^2$, которое положительно при $y_2 \neq 0$. Таким образом, условия теоремы 8 для точки

$$x_1 = 2\sqrt{\frac{S_0}{6\pi}}, \quad x_2 = \sqrt{\frac{S_0}{6\pi}}$$

выполнены. Поэтому цистерна с такими параметрами, по крайней мере, среди цистерн с близкими параметрами имеет наибольший объем.

Упражнения

Решить следующие задачи:

1. $f(x_1, x_2, x_3) = x_1 - 2x_2 + 2x_3 \to \min,$

$$x_1^2 + x_2^2 + x_3^2 = 1.$$

2. $f(x_1, x_2, x_3) = x_1 x_2 x_3 \to \min,$

$$x_1^2 + x_2^2 + x_3^2 = 1,$$
$$x_1 + x_2 + x_3 = 0.$$

3. $f(x_1, x_2) = x_1^2 + 12x_1x_2 + 2x_2^2 \to \min,$

$$4x_1^2 + x_2^2 = 25.$$

4. $f(x_1, x_2) = e^{ax_1x_2} \to \min, \qquad (a < 0)$

$$x_1^3 + x_2^3 + x_1 + x_2 - 4 = 0.$$

5. Мукомольный комбинат реализует муку двумя способами: в розницу через магазин и оптом через торговых агентов. При продаже x кг муки через магазин расходы на реализацию составляют x^2 ден. ед., а при продаже y кг муки через торговых агентов - y^2 ден. ед. В сутки для продажи выделяется 5000 кг муки. Сколько кг муки надо продавать каждым способом при минимальных затратах? Задачу решить методом Лагранжа.

Литература

1. Igor Griva, Stephen G. Nash, Ariela Sofer. Society for Industrial and Applied Mathematics, Philadelphia. 2009. (5-7, 483-486)

2. Robert M. Leekley, Applied Statiatics for Businiess and Economics, USA, 2010.

3. Alpha C. Chiang, Kevin Wainwright, Fundamental Methods of Mathematical Economics, NY 2005, (347-355)

4. Бабаджанов Ш. “Математическое программирование”. Учебное пособие. Т.: ТМИ, 2006.