В.О. Корнієнко, С.Г. Денисюк, А.А. Шиян
(Вінницький національний технічний університет)

НЕЛІНІЙНІ ЗАДАЧІ МАТЕМАТИЧНОГО ПРОГРАМУВАННЯ

Розглядається сутність нелінійного програмування, економічна нтерпретація задачі нелінійного програмування, задачі з обмеженнями-рівностями, нелінійного програмування.

Ключові слова: нелінійне програмування

Сутність нелінійного програмування

*Визначення***.** Загальна *задача нелінійного програмування* виглядає в такий спосіб:

$f(x) \to \max$ (1.1)

при

$g_i(x)=b_i$, $i \square \{1, \ldots, m_1\}$; (1.2)

$g_i(x)< b_i$, $i \square \{m_1+1,\ldots,m\}$, (1.3)

Тут $x \square R^n$; всі b_i — дійсні числа; f і g_i — функції, що приймають значення з R.

Множина всіх рішень системи рівнянь і нерівностей (1.2) - (1.3) є *множиною припустимих рішень* задачі.

Задача нелінійного програмування - це окремий випадок задач, специфіка яких полягає в тому, що X є підмножиною R^n і описується системою обмежень: рівностей (1.2) і нерівностей (1.3). Обмеження виду $g(x) \square b$ зводяться до (1.3) домноженням на (-1). Параметрами задачі нелінійного програмування є: число змінних, вигляд цільової функції - ЦФ (1.1), число, структура і праві частини обмежень.

Нехай X - множина допустимих рішень задачі (1.1) - (1.3). Всі функції g_i повинні бути визначені на X, оскільки $x \square X$ означає, що всі обмеження в точці x виконуються і, отже, ліві частини обмежень є числами. Крім того, якщо x - допустиме рішення, те модельовану операцію можна здійснити шляхом управління x і одержати деякий результат, оцінка якого, по припущенню, дорівнює $f(x)$; отже, функція f теж повинна бути визначена на X.

***Визначення*.** Функція $F(x)$ на множині $M \square R^n$ має *локальний максимум* у x_0, якщо $x_0 \in M$ й існує така околиця O точки x_0, що $F(x_0) \geq F(x)$ для всіх $x \in O \square M$.

Аналогічно визначається *локальний мінімум* функції $F(x)$ на M.

Локальний *екстремум* - це локальний максимум або локальний міні-

мум.
Локальний екстремум називають *безумовним,* якщо M□R^n**,** і умовним - у протилежному випадку.

Локальне рішення задачі нелінійного програмування - це локальний максимум функції (1.1) на множині *X,* які визначені обмеженнями (1.2), (1.3).

Очевидно, що в задачі (1.1) - (1.3) точки максимуму функції *f* на *X* збігаються із точками її найбільших локальних максимумів на цій множині. Отже, рішення задачі нелінійного програмування можна шукати серед локальних рішень цієї задачі. Тому важливо знати (знайти) необхідні та/або достатні умови локальних умовних екстремумів, які дозволяють виявити локальні рішення. Нижче будуть розглянуті такі умови для деяких класів задач.

Якщо зафіксовано множину $M \subseteq R^n$, то всі точки простору можна класифікувати в такий спосіб.

*Визначення***.** Точка *x* є *внутрішньою* для *M,* якщо якась її околиця вкладена в *M.*

*Визначення***.** Точка *x* є *зовнішньою* для *M,* якщо деяка її околиця не перетинається з *M.*

*Визначення***.** Точка *x* є *граничною* для *M,* будь-якої її околиці є і точки з *M*, і точки із $R^n \backslash M$; множина всіх граничних точок множини називають її *границею.*

Зрозуміло, що будь-яка точка із R^n по відношенню до *M* є або зовнішньою, або внутрішньою, або граничною. Множина може не мати внутрішніх точок (наприклад, пряма в R^n при $n > 1$).

Множина $X \subseteq R^n$ називається замкненою, якщо вона містить свою границю. Наприклад, коло на площині є границею диску. Внутрішні точки диску разом із колом утворять замкнуту множину.

Зауваження. З визначення граничної точки випливає, що вона є загальною межею послідовності внутрішніх точок і послідовності зовнішніх точок множини.

Рішення задачі нелінійного програмування, як правило, істотно спрощується, якщо вдається попередньо довести, що шуканий екстремум існує. Сформулюємо результати, які часто допомагають установити існування рішення.

Твердження 1. *Нехай X - множина всіх припустимих рішень задачі* (1.1) - (1.3). *Якщо X і його границя вкладені в множину M, на якій функції* g_1, ..., g_m *неперервні, то множина X замкнута.*

Доведення. Якщо *X*=0, то воно замкнуте по визначенню. Розглянемо випадок *X*□0. Нехай *x* - гранична точка *X*. Відповідно до ***зауваження*** існує послідовність точок $x_k \in X \subseteq M$ така, що $x_k \longrightarrow x$. За умовою, границя множини *X* вкладена в *M,* тому $x \in M$. Але функції g_i неперервні на *M,* тому $g_i(x_k) \longrightarrow g_i(x)$ для всіх *i*. Внаслідок властивостей межі $g_i(x) = b_i$ для $i \le m_1$ і $g_i(x) \le b_i$ при $i > m_1$. Отже, $x \in X$.

Наступна теорема Вейєрштраса дає класичну ознаку можливості розв'язання задачі оптимізації.

Теорема 1. *Нехай f*(x) *- неперервна функція на замкнутій множині* X□R^n. *Якщо множина* $Q(\alpha) = \{x \in X \mid f(x)$ □ $(\le) \alpha\}$ *не порожня і обмежена для деякого* α □ R, то *f*(x) *досягає свого максимуму (мінімуму) на X.*

Наслідок. *Якщо задача* (1.1) - (1.3) *має непусту і обмежену множину* *припустимих рішень X, функція f неперервна на цій множині, а функції g_1, ..., g_m неперервні на деякій множині, що включає X разом із її границею, то задача має рішення.*

Доведення. Якщо задача (1.1) - (1.3) задовольняє умовам, перерахованим у формулюванні наведеного вище наслідку із Теореми Веєрстштраса, то множина X її припустимих рішень замкнута за твердженням 1. Внаслідок припущення X□0, виберемо $a \in X$. Тоді $Q(\alpha)$ з теореми 1 непорожня для $\alpha = f(a)$ і обмежена (як підмножина X).

Економічна інтерпретація задачі нелінійного програмування

Багато різноманітних задач прийняття рішень допускають формалізацію у вигляді моделі нелінійного програмування. Опишемо лише одну із стандартних економічних інтерпретацій задачі нелінійного програмування: *задачу планування виробництва*.

Розглянемо операцію, метою якої є складання плану роботи економічної системи на деякий *плановий період.* Припустимо, що система може в різних сполученнях застосовувати n технологій виробництва *(технологічних способів* або просто *способів).* Помітимо, що технологія може бути не тільки виробничою, але й, наприклад, організаційною. Припустимо також, що в реалізації технологічних способів беруть участь *інгредієнти* (матеріальні ресурси, інші виробничі фактори, напівфабрикати, кінцеві продукти, корисності тощо) N_i, ..., N_m. Кожний інгредієнт може вироблятися та/або споживатися системою.

Мірою використання заданого технологічного способу в плановому періоді є його *інтенсивність,* яка, у відповідності свого визначення, може приймати лише тільки ненегативні дійсні значення (тобто додатні значення або нуль). Звичайно інтенсивність використання способу характеризують обсягом виробництва або споживання (за розглянутий період) якогось (основного для даного способу) інгредієнта. Нехай $x_j \geq 0$ — інтенсивність використання способу j ($x_j=0$ означає, що спосіб не застосовується). Тоді вектор $x=(x_j)^n{}_1$ природно назвати *планом виробництва.*

Нехай g_i(x) - *чисте споживання* системою (різниця між споживанням та виробництвом) інгредієнта N_i за плановий період при плані x; $g_i(x)<0$ означає, що в системі до кінця періоду утвориться надлишок N_i, рівний $|g_i(x)|$.

На початку періоду система має *запас* b_i інгредієнта N_i; $b_i<0$ означає, що система має зобов'язання із постачання N_i (за плановий період) у зовнішнє середовище в розмірі $|b_i|$. Допустимо, що план x дає системі (протягом планового періоду) корисність (наприклад, прибуток або чистий дохід) f(x). Тепер модель розглянутої операції можна записати в такий спосіб:

$f(x) \rightarrow \max$ (2.1)

при

$g_i(x) \leq b_i$, i □ $\{1,...,m\}$, (2.2)

$x>0.$ (2.3)

Це, звичайно, задача нелінійного програмування. Якщо b_i□ 0, то умова (2.2. 2) означає, що чисте споживання інгредієнта i у системі не може перевищувати його вихідний запас; при $b_i<0$ умова зручно записати у вигляді -g_i(x)≥|b_i|: чисте виробництво інгредієнта повинне покривати заплановані поставки.

Задачі з обмеженнями — рівностями

Попередні відомості

Визначення. Класичною задачею умовної оптимізації називають задачу

f(x) —> max при g_i(x) = b_i для i □ {1,...,m} (3.1)

Це, по суті, загальна задача нелінійного програмування (1.1) - (1.3), у якій всі обмеження є рівностями, $m_1 = m$.

Зауваження. Нерівність (1.3) еквівалентно рівнянню $g_i(x)+s^2_i=b_i$ у наступному сенсі: якщо x є рішенням нерівності, то b_i - g_i(x) ≥0, і значення $s_i=[b_i - g_i(x)]^{1/2}$ визначене, а (x,s_i) – є рішенням рівняння; якщо ж (x,s_i) є рішенням рівняння, то $b_i - g_i(x) = s^2_i$□ 0 і x — буде рішенням нерівності. В аналогічному сенсі нерівність $g(x)≥b_i$ еквівалентна рівнянню $g(x) - s^2_i= b_i$. Зокрема, x_j□0 еквівалентно $x_j=z^2_j$.

Із даного зауваження випливає, що всяку задачу нелінійного програмування можна звести до класичної задачі умовної оптимізації (втім, таке зведення не завжди доцільно робити). Припустимо, що в (3.1) функції f і g_i (i□{1,..., m}) є неперервно диференційовані на множині припустимих рішень. У цьому випадку велике теоретичне і прикладне значення має *метод множників Лагранжа,* який дозволяє, використовуючи необхідну умову локального екстремуму, знайти множину точок, які є «підозрілими на екстремум» у задачі (3.1). Нагадаємо визначення і твердження, які є необхідними для викладання цього методу. Розглянемо функцію F, яка визначена і диференційована на множині $A \subseteq \mathrm{R}^n$.

Визначення. Множина всіх рішень рівняння $F(x)=c$ (c□R) називається *поверхнею рівня* (у двомірному випадку - *лінією рівня) с* для функції F.

Поверхня рівня *проходить через точку а* (відповідно, точка *лежить* на поверхні рівня), якщо *а* задовольняє відповідному рівнянню.

Вектор, складений із частинних похідних функції F, *які* обчислені у точці x□A, називається *градієнтом* цієї функції в точці x і позначається ∇F(x) (перший символ читається «набла»), тобто

$$\nabla F(x) = \left(\frac{\partial F}{\partial x_i}(x) \right)_1^n$$

Якщо ∇F(x) = 0, то x - *стаціонарна точка* функції F.

Зауваження А. Через кожну точку $x \in A$ проходить поверхня рівня $c=F$(x) для функції F. Вектор ∇F(x) указує напрямок найшвидшого росту функції F при русі із точки x. Отже, при малих позитивних ε поверхні рівнів $c + \varepsilon$ розташовані «з боку» вектора $\nabla F(\mathrm{x})$ від поверхні рівня c.

Зауваження Б. Для порівняння із наступними результатами згадаємо також необхідну ознаку безумовного екстремуму: якщо функція F диференційована в точці x, визначена в деякій її околиці та має в x безумовний локальний екстремум, то x - стаціонарна точка функції F.

Визначення. Вектори $\boldsymbol{x}$ та $\boldsymbol{b}$ *ортогональні,* якщо $\boldsymbol{x \bullet в=0}$.

*Визначення***.** *Лінійною комбінацією* векторів $x_1,\ldots,x_k$ з R^n з коефіцієнтами $\mathrm{a}_1, \ldots, \mathrm{a}_k$ з R називається вектор $x = \sum_{i=1}^{n} a_i x_i$

Вектори $x_1,\ldots, x_k$ *лінійно залежні,* якщо якась їхня лінійна комбінація з ненульовим вектором коефіцієнтів дорівнює нулю. У протилежному випадку $x_1,\ldots, x_k$ *є лінійно незалежні.*

Ненульові лінійно залежні вектори $\boldsymbol{x}$ та $\boldsymbol{b}$ називаються *колінеарними.*

Ортогональність у двомірному і тривимірному просторах означає перпендикулярність.

Очевидно, що вектори $\boldsymbol{x}$ та $\boldsymbol{b}$ колінеарні, коли і тільки коли вони пропорційні ($в = \lambda x$, λ□R, λ□0). Ненульові колінеарні вектори $\boldsymbol{x}$ та $\boldsymbol{b}=\lambda\mathrm{x}$ *однаково спрямовані* (x□□y), якщо $\lambda > 0$, і *протилежно спрямовані* (x□□y), якщо $\lambda < 0$.

*Визначення***.** Якщо функція F *лінійна* ($F(x)=a \bullet x$, a□0), то її поверхня рівня називається *гіперплощиною.* Інакше кажучи, гіперплощина в R^n - це множина всіх рішень *лінійного рівняння* $a \bullet x=b$ при a□0.

Будь-який вектор λa при λ□0 називається *нормаллю* (до) гіперплощини $a \bullet x=b$.

Гіперплощина □$F(x_0)$□x=□$F(x_0)$□x_0 називається *дотичною гіперплощиною* до поверхні рівня функції F у точці x_0.

Зрозуміло, що гіперплощинами в R^2 і R^3 є відповідно прямі і площини.

Легко перевірити, що нормаль до гіперплощини H ортогональна до цієї гіперплощини (ортогональна до будь-якого вектору, кінці якого лежать в H). Отже, градієнт $\nabla F(\mathrm{x}_0)$ є нормаллю гіперплощини, дотичної до поверхні рівня функції F у точці x_0 і ортогональним до цієї гіперплощини.

*Визначення***.** Нехай функції F і G диференційовані в точці x_0. Поверхні $F(x)=F(x_0)$ і $G(\mathrm{x})=G(\mathrm{x}_0)$ *торкаються* в точці x_0, якщо вони мають у цій точці загальну дотичну гіперплощину, тобто $\nabla F(\mathrm{x}_0)$ і $\nabla G(\mathrm{x}_0)$ колінеарні.

Геометрична ілюстрація

Рис.1 ілюструє максимізацію функції двох змінних при одному обмеженні - рівності (n =2, m =1). Задача (3.1) у цьому випадку має вигляд

$f(x_1,x_2) \longrightarrow \max$ при $g(x_1,x_2) = b$. (a)

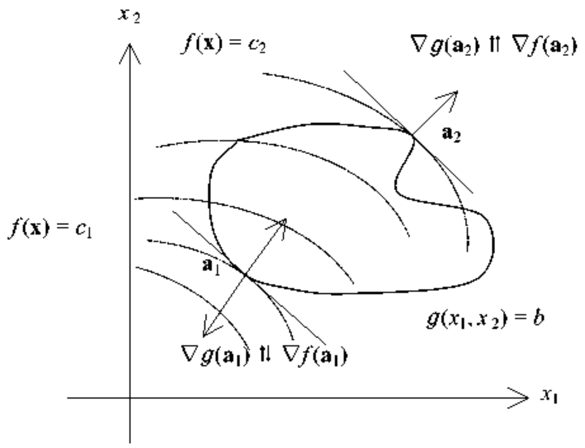

Рис. 4.3.1. Умовні екстремуми

На рисунку 4.3.1. позначені лінії рівня для функції *f*. Зазначено лінії рівнів c_1 і c_2, причому $c_1<c_2$. Множину припустимих рішень задачі зображено суцільною замкнутою кривою, яка отримується із рівняння $g(x_1,x_2)=b$ (лінія рівня b для функції g). Умовимося, що для внутрішніх точок області, обмеженої цією кривою, виконується нерівність $g(x_1,x_2)<b$ (звідси випливає, що вектор-градієнт у будь-якій точці множини припустимих рішень спрямований «назовні»).

Локальний (але не абсолютний) максимум досягається в точці a_2, тому що в її околиці немає припустимих рішень, які лежать на лініях рівнів $c>c_2$. Точка a_2 - локальне, але не оптимальне рішення задачі (а), оскільки f(x) на множині припустимих рішень (правіше й нижче точки a_2) може бути більше c_2. Локальний (і абсолютний) мінімум досягається в точці 1, тому що на лініях рівнів $c<c_1$ немає припустимих рішень.

Геометрично очевидно, що в точці локального екстремуму лінія рівня функції f повинна торкатися лінії $g(x)=b$ припустимих рішень. Тоді ці лінії в точці екстремуму повинні мати загальну дотичну, тому градієнти функцій f і g є колінеарними (лінійно залежними). На Рис.1 показані дотичні в точках a_1 і a_2; градієнти функцій f і g у цих точках не дорівнюють нулю, $\nabla f(a_1) \downarrow\uparrow \nabla g(a_1)$ і $\nabla f(a_2) \uparrow\uparrow \nabla g(a_2)$.

Розглянемо точку локального максимуму a_2. З лінійної залежності □$f(a_2)$ і □$g(a_2)$ слідує, що існують такі числа y^*_0 і y^*_1, не рівні нулю одночасно, для яких y^*_0□$f(a_2)=y^*_1$□$g(a_2)$. Інакше кажучи, вектори $y^* = (y^*_0, y^*_1)$ і a_2 є рішенням наступної системи рівнянь:

$$g(x_1,x_2)=b,$$

$$y_0\frac{\partial f}{\partial x_1}(x_1,x_2)=y_1\frac{\partial g}{\partial x_1}(x_1,x_2),$$

$$y_0\frac{\partial f}{\partial x_2}(x_1,x_2)=y_1\frac{\partial g}{\partial x_2}(x_1,x_2),$$

Серед рішень цієї системи знаходяться всі точки локальних екстремумів функції f на множині припустимих рішень а, отже, і всі рішення розглянутої задачі.

Аналогічна є необхідна умова локального рішення є справедливо в *n-вимірному* випадку.

Ознаки локального рішення.

Теорема 2 (необхідна умова першого порядку).

Нехай функції f, g_1, ..., g_m неперервно диференційовані в деякій околиці точки x^*. Якщо x^* - локальне рішення задачі (3.1), то існують число y_0 і вектор $y^* = (y_i^*)^m{}_1$, не рівні нулю одночасно й такі, що

$$y_0^* \nabla f(x^*) = \sum_{i=1}^{m} y_i^* \nabla g_i(x^*). \quad (3.2)$$

Зауваження. Рівність (3.2) разом із умовою (y_0, y^*)□$(0,0)$ ("не рівні нулю одночасно") означає лінійну залежність векторів $\nabla f, \nabla g_1, \ldots, \nabla g_m$ у точці x^*. Якщо y^*_0□0, то □$f(x^*)$ є лінійною комбінацією векторів □g_i і вектора коефіцієнтів y^*/y^*_0. Якщо ж $y_0=0$, то (3.2) виражає тільки лінійну залежність векторів □$g_i(x^*)$.

Визначення. Числа $y^*_0,...,y^*_m$, визначені в теоремі 2, називають *множниками Лагранжа* для точки x^*.

Введемо *функцію Лагранжа* для задачі (3.1):

$$L(x, y_0, y) = y_0 f(x) - \sum_{i=1}^{m} y_i [g_i(x) - b_i] \quad (3.3)$$

Легко бачити, що умови (3.2) і $g_i(x) = b_i$ відповідно еквівалентні рівностям

$$\frac{\partial L}{\partial x_j}(x, y_0, y) = 0, \quad j \in \{1,...,n\},$$

$$\frac{\partial L}{\partial y_j}(x, y_0, y) = 0.$$

Зауваження. Значення $y^*_0, ..., y^*_m$ визначені з точністю до ненульового множника: якщо набір (y_0, y^*) задовольняє умовам теореми 2, то набір $(\lambda y_0, \lambda y^*)$ при λ□0 теж задовольняє цим умовам. Тому досить розглядати тільки випадки $y^*_0=0$ і $y^*_0=1$ (якщо y^*_0□0, то можна всі множники Лагранжа розділити на y^*_0).

Отже, якщо функції f, g_1, ..., g_m неперервно диференційовані на множині припустимих рішень задачі (3.1), то всі її локальні рішення (більш того, всі локальні умовні екстремуми цільової функції) знаходяться серед рішень наступної системи рівнянь при $y_0 \in \{0,1\}$:

$$\frac{\partial L}{\partial x_j}(x, y_0, y) = y_0 \frac{\partial f}{\partial x_j}(x) - \sum_{i=1}^{m} y_i \frac{\partial g_i}{\partial x_j}(x) = 0,\ 1 \le \mathrm{j} \le n; \quad (3.4)$$

$$\frac{\partial L}{\partial y_j}(x, y_0, y) = g_i(x) - b_i = 0,\ 1 \le i \le m. \quad (3.5)$$

Підкреслимо, що частинна похідна функції Лагранжа по y_0 не бере участь у рівняннях (3.5). Отже, множина «підозрілих на екстремум» припустимих точок задачі (3.1), яка включає всі її рішення, складається зі стаціонарних точок функцій *L(x,0,y)* і *L(x,1,y)* (порівняйте із результатом, сформульованим у наведеному вище зауваженні) і точок, у яких не всі функції f, g_1, ..., g_m неперервно диференційовані.

*Визначення***.** Якщо для набору (x,y_0,y) при y_0□0 виконані умови (3.4) і (3.5), то *x називають регулярною* точкою задачі (3.1).

Наступне очевидне (дивись одне із наведених вище зауважень) твердження дає найбільш уживану достатню *умову регулярності.*

Твердження 2. *Якщо точка x є припустимою в задачі (3.1) і задовольняє необхідним умовам першого порядку (дивись Теорему 2), а градієнти функцій g_i у цій точці є лінійно незалежними, то вона є регулярною точкою задачі.*

Сформулюємо тепер необхідну умову другого порядку, яку використовують для виключення тих стаціонарних крапок функції Лагранжа, які не є локальними рішеннями класичної задачі. Попередньо помітимо, що

$$\frac{\partial^2 L}{\partial x_j \partial x_k}(x, y_0, y) = y_0 \frac{\partial^2 f}{\partial x_j \partial x_k}(x) - \sum_{i=1}^{m} y_i \frac{\partial^2 g_i}{\partial x_j \partial x_k}(x) \quad (3.6)$$

Теорема 3 (необхідна умова другого порядку).

Нехай функції f, $g_1,\ldots,g_m$ є двічі диференційовані в регулярній точці x і неперервно диференційовані в деякій її околиці. Якщо x — локальне рішення задачі (3.1), то

$$\sum_{j=1}^{n}\sum_{k=1}^{n} \frac{\partial^2 L}{\partial x_j \partial x_k}(x,1,y) \cdot h_j \cdot h_k \le 0 \quad (3.7)$$

для всякого вектора h□R^n*, що задовольняє умовам*

$$\nabla g_i(x) \cdot \overline{h} = \sum_{j=1}^{n} \frac{\partial g_i}{\partial x_j}(x) \cdot h_j = 0, \quad i \in \{1,\ldots m\} \quad (3.8)$$

і кожного ***y****, що задовольняє умовам (3.4) при* $y_0=1$.

Приведемо також одну із достатніх умов локального рішення задачі (3.1).

Теорема 4 (достатня умова другого порядку).

Нехай функції f, $g_1, \ldots, g_m$ *двічі диференційовані в припустимій точці x. Якщо існують не рівні нулю одночасно число* y_0 *і вектор* ***y****, які задовольняють умовам (3. 4), і*

$$\sum_{j=1}^{n}\sum_{k=1}^{n}\frac{\partial^2 L}{\partial x_j \partial x_k}(x, y_0, y)\cdot h_j \cdot h_k \le 0 \quad (3.9)$$

для всякого ненульового вектора h□R^n*, що задовольняє умовам (3.8), то x є локальним рішення задачі (3.1).*

Зауваження. Умова (3.8) означає, що h - *дотичний напрямок:* він є ортогональним до градієнтів функцій g_i і є, отже, дотичним у точці x до множини припустимих рішень. У деякому сенсі це і є *припустимий напрямок:* при нескінченно малому зрушенні із x у напрямку h ми «потрапимо» у припустиму точку. Умови (3.7) і (3.9) разом із необхідними умовами (3.4) гарантують не зростання цільової функції при такому зрушенні.

При кожному значенні y_0□$\{0,1\}$ система (3.4), (3.5) містить $n+m$ невідомих $(x_1,\ldots,x_{,,}, y_1,\ldots,y_m)$, стільки ж обмежень і, в принципі, може бути вирішена. Визначивши в такий спосіб множину «підозрілих» точок, потрібно вибрати з-посеред них рішення задачі. Якщо можливість розв'язання задачі доведена, то відбір здійснюють прямим обчисленням функції f у підозрілих точках (у шуканих точках значення функції максимальні). У протилежному випадку доводиться використовувати необхідні умови другого порядку і достатні умови локального рішення. Це і є метод множників Лагранжа для класичної задачі умовної оптимізації.

Теореми 3 і 4 дають «майже необхідну і достатню» умову того, що регулярна точка є локальним рішенням задачі (3.1): різниця – лише у знаках нерівностей (3.7) і (3.9). Якщо в точці x умова (3.7) виконується як рівність для деяких h, що задовольняють (3.8), то потрібні більш сильні засоби, щоб з'ясувати, чи є x локальним рішенням. Зауважимо також, що застосування теорем 3 і 4 не є простою задачею.

Однак для нерегулярних підозрілих точок ситуація ще гірша: аналіз таких точок вимагає, як правило, ретельного урахування специфіки задачі.

Приклад 1 (схема застосування методу множників Лагранжа).

Розглянемо задачу

$f(\mathrm{x}) = x_1^2 + x_2^2 + x_3^2 \rightarrow \min$ при $4x_1 + x_2^2 + 2x_3 = 14$. (а)

Тут множина припустимих рішень $X=\{x$□$R^3 \mid 4x_1 + x_2^2 + 2x_3 = 14\}$. Помітимо насамперед, що рішення задачі існує. Дійсно, нехай $x_0=(0,0,7)$; $f(x_0) = 49$ і x_0 задовольняє обмеженню задачі. Множина $Q=\{x$□$X \mid f(x) < 49\}$ непорожня (x_0□Q) і обмежена (з $f(x)$□49 випливає, що $||x||$□7), тому $f(x)$ має на X точку мінімуму внаслідок умов *Твердження 1* і *Теореми 1*. Перейдемо до еквівалентної задачі максимізації:

$$-x^2_1 - x^2_2 - x^2_3 \longrightarrow \max \quad \text{при} \quad 4x_1 + x^2_2 + 2x_3 = 14. \qquad \text{(b)}$$

Будуємо функцію Лагранжа:

$$L(x, y_0, y_1) = -y_0(x^2_1 + x^2_2 + x^2_3) - y_1(4x1 + x^2_2 + 2x_3 - 14)$$

і складаємо систему рівнянь (3.4), (3.5):

$$\partial L/\partial x_1 = -2y_0x_1 - 4y_1 = 0, \qquad \text{(c)}$$
$$\partial L/\partial x_2 = -2y_0x_2 - 2y_1x_2 = 0, \qquad \text{(d)}$$
$$\partial L/\partial x_3 = -2y_0x_3 - 2y_1 = 0, \qquad \text{(e)}$$
$$\partial L/\partial y_1 = -(4x_1 + x_2 + 2x_3 - 14) = 0, \qquad \text{(f)}$$

Тут $(y_0, y_1) \neq (0, 0)$ і $y_0 \in \{0,1\}$. При $y_0=0$ з (e) одержуємо $y_1=0$, що неможливо; тому $y_0=1$. Тоді з (c), (e) і (d) слідує, що $x_1= -2y_1$, $x_3= -y_1$ *і* $(y_1 + 1)x_2=0$ відповідно. Остання рівність дає $x_2=0$ або $y_1= -1$.

Якщо $x_2=0$, то підстановка x_1, x_2 *і* x_3 в (f) дає $10y_1= -14$, $y_1= -7/5$, звідки знаходимо рішення системи: $x_1=(14/5,0,7/5)$. Якщо $y_1= -1$, то $x_1=2$ і $x_3=1$; підставляючи ці значення в (f), знайдемо $(x_2)^2 = 4$. Одержали, таким чином, ще дві підозрілі точки: $x_2=(2, 2,1)$ і $x_3=(2, -2,1)$.

Обчислимо значення цільової функції задачі (а) у знайдених підозрілих точках: $f(x_1)=49/5$, $f(x_2)=f(x_3)=9$. Точка x_1, отже, не може бути точкою мінімуму в задачі (а). Але задача, як показано вище, має рішення, і воно знаходиться серед точок x_2 і x_3 по Теоремі 2. Отже, x_2 і x_3 — шукані точки мінімуму, що дорівнює 9.

Помітимо, що наведене рішення має лише методичну цінність: якщо виразити x^2_2 з обмеження задачі і результат підставити в цільову функцію, то одержимо нескладну задачу безумовної оптимізації.

Продовжуючи приклад, розглянемо задачу максимізації тієї ж функції при тому ж самому обмеженні:

$$f(x) = x^2_1 + x_2 + x_3 \longrightarrow \max \quad \text{при} \quad 4x_1 + x_2 + 2x_3 = 14. \qquad \text{(g)}$$

Нехай $x(\varepsilon) = (\varepsilon/2, 0, 7 - \varepsilon)$. Точка $x(\varepsilon)$ припустима в задачі (g) при будь-якому $\varepsilon \in R$. Але $f(x(\varepsilon)) = (5/4)\varepsilon^2 - 14\varepsilon + 49 \longrightarrow \infty$ при $\varepsilon \longrightarrow -\infty$, отже, задача (g) не має рішення. Зробимо, однак, вигляд, що ми цього не знаємо, і застосуємо необхідну умову другого порядку.

Функція Лагранжа для задачі (*g*) має вигляд

$$L(x, y_0, y_1) = y_0(x_1 + x_2 + x^2_3) - y_1(4x_1 + x^2_2 + 2x_3 - 14).$$

Система рівнянь для визначення підозрілих точок відрізняється від системи (c) - (f) тільки відсутністю мінусів у перших доданках перших трьох рівнянь і має ті ж рішення: $x_1=(14/5,0,7/5)$, $x_2=(2, 2, 1)$ і $x_3=(2,-2,1)$; $y_0=1$

для всіх рішень, y_1=7/5 для x_1, $y_1=1$ для x_2 і x_3. Порівняння значень цільової функції показує, що x_2 і x_3 не можуть бути точками максимуму.

Точка x_1 регулярна в задачі (g). Нерівність (3.7) приймає вигляд

$$2h^2_1 + 2(1 - y_1)h^2_2 + 2h^2_3 < 0,$$

причому y_1 = 7/5. Тобто,

$$2h^2_1 -(4/5)\, h^2_2 + 2h^2_3 < 0. \quad \text{(h)}$$

Якщо x_1 - рішення задачі (g), то нерівність (h) повинна виконуватися при всіх h□R^3 таких, що $4h_1 + 2x^1_2h_2 + 2h_3=0$ (умова (3.8) теореми 3), де x^1_2=0 - другий компонент вектора *1*. На вектор *h*, отже, накладена умова

$$2h_1 + h_3=0. \qquad \text{(i)}$$

При $h=(1,0,-2)$ умова (i) виконана, а нерівність (h) порушується. Отже, *1* не вирішує задачу максимізації (по теоремі 3).

Отже, жодне локальне рішення не є оптимальним, задача (g) не має рішень, функція *f* на *X* необмежена зверху.

Загальна задача нелінійного програмування

Розглядаємо задачу (1.1) - (1.3). Можливі, звичайно, випадки m_1=0 (немає обмежень - рівностей) і $m_1=m$ (немає обмежень - нерівностей). В останньому випадку ми одержуємо класичну задачу умовної оптимізації, про яку йшла мова в розділі 3. Нехай *X* - множина припустимих рішень задачі.

Визначення. Для всякого x□ X введемо множину (номерів) обмежень, *активних* (які виконуються як рівності) у точці x: $I(x) = \{i \mid g_i(x) = b_i\}$.

Зрозуміло, що $\{1,...,m_i\}$ □ $I(x)$ для всіх $\mathbf{x}$□X.

Необхідні умови першого порядку.

Їх дає наступна теорема Каруша—Джона (порівняйте із Теоремою 2).

Теорема 5 (необхідні умови Ф. Джона). *Припустимо, що: функції* g_i *при* i□$I(x^*)$ *диференційовані в точці* x^*; *функції* f *і* g_i *при* i□$I(x^*)$, $i > m_1$ *диференційовані, а функції* g_i *при* i□ m_1 *неперервно диференційовані в деякій околиці цієї точки. Тоді, якщо* x^* — *локальне рішення задачі (1.1) - (1.3), то існують число* y^*_0□0 *і вектор* $y^*=(y_i^*)^m_1$, *не рівні нулю одночасно, які задовольняючим наступним умовам:*

$$y^*_0\, \nabla f(x^*) - \sum_{i=1}^{m} y^*_i \cdot g_i(x^*) = 0, \qquad (4.1)$$

$$y^*_i\,\big(g_i(x^*) - b_i\big) = 0, \quad \mathrm{i} > \mathrm{m}_1, \qquad (4.2)$$

$$y^*_i \ge 0, \quad i > m_1. \qquad (4.3)$$

Зауваження. У задачі (1.1) - (1.3) «підозрілими на екстремум» є припустимі точки, для яких або виконуються умови (1.2), (1.3) і (4.1) - (4.3), або ж порушені умови гладкості, накладені на функції $f, g_1, ..., g_m$ Теоремою 5. При цьому значення $y^*_0, ...,y^*_m$ визначені з точністю до позитивного множника: якщо набір (y^*_0,y^*) задовольняє умовам Теореми 5, то набір $(\lambda y^*_0,\lambda в^*)$ при $\lambda>0$ теж задовольняє цим умовам. Тому можна вважати, що y^*_0□$\{0,1\}$ (порівняйте із одним з наведених вище зауважень).

Таким чином, при y^*_0□$\{0,1\}$ для визначення значень $n + m$ змінних $x_1, ..., x_n, y_1, ..., y_m$ маємо $n + m$ рівнянь: n рівнянь

$$y^0 \frac{\partial f}{\partial x_j}(x) = \sum_i y_i \frac{\partial g_i}{\partial x_j}(x), \quad j \in \{1,...n\} \qquad (4.4)$$

Вони описані векторною рівністю (4.1). Маємо також m_1 рівнянь (1.2) і $m - m_1$ рівнянь (4.2); крім того, y□0 при $y_0=0$ і $(y_{mi+l},...,y_m)$□0 по (4.3).

Числа $y^*_0, y^*_1, ...,y^*_m$ - це *множники Лагранжа* для задачі (1.1) - (1.3). Підкреслимо, що Теорема 5 не обмежує за знаком множники Лагранжа, які відповідають обмеженням - рівностям. Рівності (4.2) називають *умовами доповнюючої нежорсткості*; вони вимагають, щоб множники Лагранжа, які відповідають неактивним обмеженням, дорівнювали нулю (або, що еквівалентно, щоб обмеження, які відповідають позитивним множникам Лагранжа, були активними).

Зауваження. З урахуванням (1.3) і (4.3), умови доповнюючої не жорсткості можна сформулювати так: *якщо* $i>m_1$, *то* $g_i(x^*) - b_i \le 0$ з рівністю *при* $y^*_i>0$. Якщо доповнити (4.2) вимогою позитивності множників Лагранжа для активних обмежень - нерівностей (з умови $g_i(x^*)=b_i$ *випливає* $y^*_i>0$ при $i>m_1$), то одержимо *умови суворої доповнюючої не жорсткості.* Вони вима- гають, щоб у добутку $y^*_i(g_i(x^*) - b_i)$ для $i>m_1$ один і тільки один співмножник дорівнював нулю.

При $y^*_0=0$ точка x^* не містить інформації про градієнт цільової функції, а перевірка оптимальності таких точок істотно утруднена. Це пояснює інтерес до регулярних точок, які вище були визначені для класичної задачі умовної оптимізації, а для загальної задачі нелінійного програмування визначаються в такий спосіб.

*Визначення****.*** Якщо набір (x,y_0,y) при $y_0>0$ задовольняє умовам (1.2), (1.3) і (4.1) - (4.3), то x - *регулярна* точка задачі (1.1)- (1.3).

Достатня умова регулярності, сформульована у Твердженні 2, у загальному випадку перетворюється у вимогу лінійної незалежності градієнтів активних обмежень. Наступний результат показує, що ця умова забезпечує також і одиничність вектора множників Лагранжа (з точністю до позитивного множника).

Теорема 6 (достатня умова регулярності). *Нехай* x^*□X, *вектори* □$g_i(x^*)$ *для* i□$I(x^*)$ *(градієнти активних у точці* x^* *обмежень) лінійно незалежні та* *існують y^*_0□0 і y^*, не рівні нулю одночасно і такі, що набір (x^*, y^*_0 , y^*) задо- вольняє умовам (4.1) -(4.3). Тоді $y^*_0>0$, x^* - регулярна точка задачі (1.1) - (1.3), а всі вектори множників Лагранжа для точки x^* пропорційні (y^*_0, y^*) і мають вигляд $c\bullet(y^*_0, y^*)$, $c>0$.*

Для регулярного локального рішення задачі нелінійного програмуван-
ня (при $y_0=1$) спрощується умова (4.1):

$$\nabla f(x^*) = \sum_{i=1}^{m} y^*_i \cdot \nabla g_i(x^*), \tag{4.5}$$

При цьому праворуч, внаслідок (4.2), можна обмежитися підсумову- ванням по i□$I(x^*)$. У такій модифікації необхідні умови, сформульовані в Те- оремі 5, називають *умовами Куна - Такера.*

Зауваження. З Теореми 6 випливає, що для регулярної точки, у якій градієнти активних обмежень лінійно незалежні, існує єдиний вектор множ- ників Лагранжа з $y^*_0=1$, що задовольняє умовам Куна—Такера.

Для задачі (1.1) - (1.3) визначимо функцію Лагранжа $L(x,y_0,y)$ у такий же спосіб, як для класичної задачі умовної оптимізації - дивись (3.3). Тоді (4.1) і (4.5) можна записати у вигляді

$$\frac{\partial L(x^*, y^*_0, y^*)}{\partial x_j}(x^*) = 0$$ і

$$\frac{\partial L(x^*, 1, y^*)}{\partial x_j}(x^*) = 0, \quad 1 \le j \le n.$$

Література

1. Губко М.В., Новиков Д.А. Теория игр в управлении организационными системами. М.:ИПУ, 2005. – 138с.
2. Печерский С.Л., Беляева А.А. Теория игр для экономистов. Вводныйкурс. Учебное пособие. – СПб.: 2001. – 342 с.
3. Данилов В.И. Лекции по теории игр /КЛ/ 2002/ 001. – М.: Российскаяэко- номическая школа, 2002
4. Вилкас Э.Й. Аксиоматическое определение значения матричной игры // Теория вероятностей и ее применения. 1962. Том. 8. № 3. С.324-327.
5. Нейман Д., Моргенштерн О. Теория игр и экономическое поведение.М.: Наука, 1970.
6. Петросян Л.А., Зенкевич Н.А., Семина Е.А. Теория игр. М.:

Высшая школа, 1998.
7. Мулен Э. Кооперативное принятие решений: аксиомы и модели. М.: Мир, 1991.
8. Lave Ch., March J.G. An Introduction to Models in Social Sciences. N.Y.,1978.
9. Persson T., Tabellini G. Political Economics: Explaining Economic Policy.- Cambridge, MA: MIT Press, 2000. – 533 p.
10. Яковлев И.Г. Информационно-аналитические технологии и политическое консультирование // Полис. №.2. - 1998.
11. Саати Т. Математические методы исследования операций. М., 1962.
12. Black Duncan On the Rationale of Group Decision-making. – The Journal of Political Economy, 56, P. 23-34.
13. Downs Antony An Economic Theory of Democracy, New York; Harper & Row.
14. Hotelling, Harold Stability in Competition // Economic Journal/ - № 39.
15. Шиян А.А. Управление государством. Учебник для Президентов, Премьеров и Парламентариев. - Винница,2005. – Веб-сайт http://soctech.narod.ru.
16. Шиян А.А. Формирование иерархических социальных структур как способ проведения избирательных кампаний. Теория и результаты апробации. Анализ одной избирательной кампании по одномандатному избирательному округу на выборах в государственную Думу России в 1999 году // Политический маркетинг (Москва, Россия). - 2000. - №3. - с.9-43.
17. Шиян А.А. Выборы в Украине: технологический тупик // Политический маркетинг (Москва). – 2006. - №5. – С.31-38.