Тема 17

Държавен изпит

специалност

Приложна математика

Задача на линейното оптимиране. Основни теореми

Анотация

Задача на линейното оптимиране. Основни теореми

1. Задачана линейното оптимиране в общ вид. Канонична форма. Канонично многостенно множество. Теорема за представяне на канонично многостенно множество.

2. Основни теореми на линейното оптимиране за канонична линейна задача.

1.

Общата задача на линейното оптимиране (ЛО) е частен случай на задача на оптимирането – да се минимизира/максимизира функция при дадени условия. В случая на ЛО функцията, която искаме да минимизираме е линейна по всички променливи и също условията (ограниченията) са линейни уравнения/неравенства. Да запишем формално:

$$(L)\left|\begin{array}{l} \min/\max \quad z(x) = c_0 + \sum_{j=1}^{k} c_j x_j \\ \sum_{j=1}^{k} a_{i,j} x_j \overset{\leq}{\underset{\geq}{=}} b_i \\ i = 1, 2, 3, \dots \\ b_i \in \Re, a_{i,j} \in \Re, c_i \in \Re \end{array}\right.$$

Понякога, за удобство, неравенствата за неотрицателност на променливите, т.е. $x_j \geq 0, j \epsilon J \subset \{1,2,\dots,k\}$ се записват отделно. За да е по- нагледен записа, ще запишем (L) матрично (векторите са вектор- стълбове):

$$(L)\left|\begin{array}{l} \min/\max \quad z(x) = c_0 + c^T.x \\ a_i^T.x \overset{\leq}{\underset{\geq}{=}} b_i \quad , i = 1, 2, \dots, m \\ x_j \geq 0, j \grave{o} J \subset \{1, 2, \dots, k\} \end{array}\right.$$

Да забележим, че решаването на (L) не зависи съществено от c_0 – то се добавя към функцията без да зависи от променливите. Поради тази причина, от тук нататък ще го пропускаме.

Дефиниция 1.1 Точка от R^k, за която са изпълнени всички условия от (L) се нарича допустима. В такъв смисъл, решение на (L) са тези допустими точки, за които се достига min/max на z.

За да работим удобно със задачите на ЛО ги превръщаме в задачи в „каноничен“ вид.

Дефиниция 1.2 Канонична задача е задача от вида (L) , с критерий “min” , всички огранчения на която са равенства с неотрицателни десни страни и върху променливите на която е наложено условие за неотрицателност.

Правила за свеждане на задача от тип (L) към канонична:

Нека имаме задача (L);

- Ако условието в задачата е за max , тогава го сменяме на min като обръщаме знака на функцията, т.е. $max\, c^T x\; min - c^T x$

- Тези променливи, върху които няма ограничение за неотрицателност се заменят в целевата функция и ограниченията с разликата на две неотрицателни променливи, т.е. $-¿\geq 0$ $+¿\geq 0, x_j^{¿}$ $-¿, x_j^{¿}$ $+¿-x_j^{¿}$ $x_j = x_j^{¿}$
- Неравенствата от вида $a_i^T.x \leq b_i$ се заменят с равенства чрез въвеждане на нова неотрицателна променлива $x_{k+i} = b_i - a_i^T.x \geq 0$, т.е. след въвеждането имаме равенството $a_i^T.x + x_{k+i} = b_i$
- Неравенствата от вида $a_i^T.x \geq b_i$ се заменят с равенства чрез въвеждане на нова неотрицателна променлива $x_{k+i} = a_i^T.x - b_i \geq 0$, т.е. след въвеждането имаме равенството $a_i^T.x - x_{k+i} = b_i$
- Ако някоя дясна страна е отрицателен знак, умножаваме по (-1)

Твърдение. Изходната задача (L) и получената(съответна) задача (K) са еквивалентни в смисъл, че на всяка допустима точка x_L по единствен начин се съпоставя допустима точка x_K, така че стойностите на целевата функция може да се различават евентуално по знак.

Тъй като в каноничната задача (K) има само ограничения от тип равенства, то тя може да се запише още по- удобно:

$$(K)\left|\begin{array}{l} \min/\max \quad z(x) = c^T.x \\ Ax = b \\ x \geq 0 \end{array}\right.$$

, $A \epsilon R^{mxn}, x \epsilon R^n, c \epsilon R^n, b \epsilon R^m$, където вектор- редовете на A отговарят на ограниченията, а вектор- стълбовете – на променливите.

Дефиниция 1.3 Множеството от точки, за които се удовлетворяват условията в (К) се нарича канонично многостенно множество (КММ). Това е така, понеже КММ е сечение на хиперравнини и части от хиперравнини в R^n .

Характеризация на КММ

Дефиниция 1.4 Нека x и y са различни точки от R^n . Тогава всички точки от вида $\lambda x+(1-\lambda)y, \lambda \grave{o}[0,1]$ се наричат изпъкнала комбинация на x и y (това са точките от отсечката [x,y]) Вътрешните точки от тази отсечка $\lambda \epsilon (0,1)$ се наричат истинска изпъкнала комбинация на x и y.

Дефиниция 1.5 Казваме, че едно подмножество C на R^n е изпъкнало, ако заедно със всеки две негови точки то съдържа и отсечката между тях.

Дефиниция 1.6 Казваме, че една точка от изпъкнало множество C е връх, ако не може да се представи като истинска изпъкнала комбинация от кои да е две точки от C.
Пример: върховете на изпъкнал многоъгълник са върхове, всяка точка от контура на кръг също е връх.

Дефиниция 1.7 изпъкнала комбинация на k точки $x_1,..,x_k$ от R^n наричаме всяка точка от вида $x=\sum_{i=1}^{n}\lambda_i x_i, \lambda_i \geq 0, \sum_{i=1}^{n}\lambda_i=1$

Дефиниция 1.8 Изпъкнала обвивка на k точки $x_1,..,x_k$ от R^n наричаме множеството от всички тяхни изпъкнали комбинации, т.е.

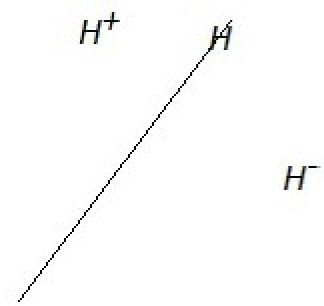
$$\{x \epsilon R^n \vee x=\sum_{i=1}^{n}\lambda_i x_i, \lambda_i \geq 0, \sum_{i=1}^{n}\lambda_i=1\}$$

Дефиниция 1.9 k- мерен симплекс е изпъкналата обвивка на k+1 точки $x_0,..,x_k$, за които векторите $x_1-x_0, x_2-x_0,..,x_k-x_0$ са ЛНЗ . $x_0,..,x_k$ са върхове на този симплекс.
Пример: При k=1 – отсечка, при k=2 – триъгълник, при k=3 – тетраедър.

Тема 17

Твърдение: Сечението на изпъкнали множества е изпъкнало множество.
Доказателство: Наистина, нека имаме k изпъкнали множества и вземем две точки x и y от сечението им. Понеже те принадлежат на всяко едно от тях, а те са изпъкнали, то и всяка изпъкнала комбинация на x и y лежи във всяко едно от множествата. Но следователно лежи и във сечението им.

Дефиниция 1.10 Подмножество на R^n от вида $H=\{x\epsilon R^n | a^T x=b\}$ е хиперравнина с нормален вектор a (ортогонален на всеки вектор от хиперравнината). Множеството от решенията на $a^T x\leq b$ се нарича затворено полупространство и се бележи с H^+. H^- пък е множеството от решенията на $a^T x\geq b$. Ясно е, че всяка хиперравнина разделя R^n на две затворени полупространства със сечение самата хиперравнина.



Твърдение: Хиперравнините и затворените хиперравнини са изпъкнали множества.
Следствие: КММ е изпъкнало.

Лема 1.1
Нека x е в M. Ако вектор стълбовете на A, съответни на ненулевите координати на x са ЛЗ, то съществува ненулев вектор z, за който:
- Az=0
- z_i=0 ако x_i=0

$+¿$
$-¿;t^{¿}$
$t^{¿}:x+tz\epsilon M$
$+¿:\forall t\epsilon¿$
- $-¿<0<t^{¿}$
$+¿:t^{¿}$
$-¿,t^{¿}$
$\exists t^{¿}$

Теорема 1.1 (алгебрична характеризация на КММ – НДУ за връх в КММ)
Нека е дадено КММ M. Точка x от M е връх на M тогава и само тогава когато вектор-стълбовете на матрицата A, съответстващи на положителните координати на x са ЛНЗ.
Текста, който следва в червено не е задължителен за темата.

Нека уточним, че можем да считаме, че матрицата A има максимален ранг и $m\leq n.$ Това е така, понеже ако A има линейно зависими редове, то или задачата е несъвместима (ако десните страни на ограниченията на ЛЗ редове са несъвместими) или може да се опрости чрез задраскване на редове до получаване на пълен ранг. Освен това, наистина ако при пълен ранг имаме m>n , то системата ще е преопределена и несъвместима.

Важно понятие в ЛО е базистно допустимо решение. Понеже имаме пълен ранг, то можем да намерим в матрицата A m ЛНЗ стълба, отговарящи на индекси от множеството B . Останалото множество от индекси бележим с $N.$ Можем да пренаредим така променливите, че матрицата A да се представя във вида $A=(B\vee N)$, където B отговаря на B , а N - на N . Тогава, всеки вектор x от R^n можем да го разбием $x=\binom{x_B}{x_N}$. Тогава, за всяко допустимо решение имаме $(B\vee N)\binom{x_B}{x_N}=Bx_B+Nx_N=b$. Поради избора на множеството B , матрицата B е обратима. Имаме $x_B=B^{-1}b-B^{-1}Nx_N$. По този начин получихме, че стойностите на x_N (n-m на брой) напълно определят x_B . В специалния случай, когато $x_N=0$ да означим $\overset{¿}{x_B}=B^{-1}b$. На всяка такава матрица B отговаря по едно $\overset{¿}{x_B}$.Наричаме ($\overset{¿}{x_B}$,0) базистно решение. Ако някое базистно решение е и допустимо, т.е. $B^{-1}b\geq 0$, то се нарича **базистно допустимо решение(БДР).** Последното е важно поради

Теорема 1.2

Нека е дадено M – КММ. Тогава всяко x от M следните са еквивалентни:

- x е БДР
- x е връх

Дефиниция 1.11 БДР с положителни базистни координати се нарича неизродено. В противен случай, т.е. сред $\overset{¿}{x_B}$ има поне една нула, е изродено. Съответно един връх е изроден ако съответното му БДР е изродено и обратното.

Да отбележим, че когато един връх е неизроден, на него съответства единствена базистна матрица B. Ако е изроден, му съответства множество от различни базистни матрици.

Теорема 1.3 Всяко КММ има краен брой върхове.

Доказателство: От теорема 1.2 следва, че върховете са най- много колкото са базистните матрици. Те от своя страна са по- малко от $\binom{n}{m}$

Дефиниция 1.12 (посока)
Нека S е множество в R^n и d е ненулев вектор от това пространство. d се нарича посока в S , ако за всяка точка $x_0 \epsilon S$ всички точки от лъча $\{x_0 + td \vee t \geq 0\}$ се съдържат в S.
Ясно е , че ако едно множество има посока, то е неограничено в R^n.

Теорема 1.4 (характеризация на посоките в КММ)
Един вектор d в R^n е посока в КММ М, тогава и само тогава когато $Ad = o \wedge d \geq 0$.

Теорема 1.5 (теорема за представяне на КММ)
Нека е дадено КММ М и нека множеството му от върхове е V, $V = \{x_i \vee i \epsilon I\}$. Тогава всяка точка х от М може да се представи във вида

$$(1)\; x = \sum_{i \epsilon I} \lambda_i x_i + d, \lambda_i \geq 0, \sum_{i \epsilon I} \lambda_i = 1,$$

където d е посока или d е нулевият вектор.
Доказателство: Индукция по броя на „+“ координат на точки в М.
Нека минималният брой положителни координати, които има всяка точка в М е p_0.

Първа стъпка от индукцията: представянето е в сила за точките от М с точно p_0 координати.

1сл. p_0=0. В този случай нулевия вектор е връх в М – за него очевидно представянето е в сила.

2сл. p_0>0. Нека х е точка с точно p_0 координати $x = x(x_1, x_2, .., x_{p_0}, 0,..,0)$ (такова представяне постигаме след евентуално пренареждане/преименуване на променливите). Ще докажем, че х е връх. Да допуснем, че не е връх. Тогава стълбовете в А : $A_1, A_2, \ldots, A_{p_0}$ са ЛЗ и по Лема следва, че съществува ненулев вектор z, такъв, че $z_i = 0 \; при \; x_i = 0, Az = 0, \exists t ¿ \; -¿ < 0, \exists t ¿ \; +¿ > 0$ и $x + tz \,|\, +¿ \, ¿$ се съдържа в М. При това поне едно от $t ¿ \, -¿, t ¿ \, +¿$ е крайно. Нека за определеност е $t ¿ \, -¿$.Но за него знаем, че

$$-¿ := max\{-\frac{x_j}{z_j} \vee z_j > 0\} = \frac{-x_{j_0}}{z_{j_0}}, j_0 \epsilon \{1,2,..,p_0\}$$
$$t^{¿}$$

Тогава за j$_0$-та координата на $\overset{-¿\,z}{x+t^{¿}}$ имаме че е 0. Но това е противоречие, тъй като тази точка е от M и същевременно има най- много p_0-1 ненулеви координати.

Втора стъпка от индукцията: индукционно предположене: Нека (1) е в сила за точките от M с по- малко от p положителни координати (тук p е поне p$_0$+1).

Трета стъпка от индукцията: Да вземем една точка от M с точно p положителни координати, $x=x(x_1, x_2, .., x_p, 0,..,0)$

Ако x е връх, теоремата е доказана. Нека x не е връх. Тогава вектор стълбовете $A_1, A_2, \ldots, A_p$ са ЛЗ и съществува ненулев вектор z, такъв, че

$$+¿>0$$
$$-¿<0, \exists t^{¿}$$
$$z_i=0 \; при \; x_i=0, A\,z=0, \exists t^{¿}$$

и $x+tz$ | $+¿$ $¿$ се съдържа в M.

Имаме три възможности за знаците на z, които отговарят на възможностите за това $+¿$ $-¿, t^{¿}$ $t^{¿}$ да са крайни.

1. z има исамо “+” и „-„ координати. Тук $\overset{-¿}{t^{¿}}$ е крайно, $\overset{+¿}{t^{¿}}$ е крайно

2. z има само “+” и нулеви координати. Тук $\overset{-¿}{t^{¿}}$ е крайно, $\overset{+¿=+\infty}{t^{¿}}$.

3. z има само “-” и нулеви координати. Тук $\overset{+¿}{t^{¿}}$ е крайно, $\overset{-¿=-\infty}{t^{¿}}$.

Разглеждаме тези случаи последователно:

1. Разглеждаме точките $+¿\,z$ $-¿\,z, x''=x+t^{¿}$ $x'=x+t^{¿}$, които са крайни. Разсъждавайки какво в края на първа стъпка от индункцията стигаме до извода, че x', x'' имат представяне (1). Имаме

$$x' = \sum_{i\epsilon I} \lambda_i' x_i + d' ; x'' = \sum_{i\epsilon I} \lambda_i'' x_i + d''$$

Но x е вътрешна точка за отсечката [$x', x''¿$ и следователно има положително число между 0 и 1 , λ , за което $x = \lambda x' + (1-\lambda) x''$. То дори може да се намери точно:

$$-¿ zt$$
$$+¿ z$$
$$+¿$$
$$-¿ + (1-\lambda) t^{¿} z$$
$$\lambda t^{¿}$$
$$x + t^{¿} = x + ¿$$
$$x + t^{¿} + (1-\lambda) ¿$$
$$x = \lambda ¿$$

От тук имаме, че $\begin{matrix} +¿ - t^{-¿} \\ t^{¿} \\ \frac{-t^{-¿}}{¿} \\ +¿ = 0, \text{т.е.} \lambda = ¿ \\ -¿ + (1-\lambda) t^{¿} \\ \lambda t^{¿} \end{matrix}$. И така,

$$x = \lambda x' + (1-\lambda) x'' = \sum_{i\epsilon I} (\lambda \lambda_i' + (1-\lambda) \lambda_i'') x_i + \lambda d' + (1-\lambda) d''$$

Проверява се лесно, че това е представяне от вида (1).

2. Тук $\begin{matrix} -¿ \\ t^{¿} \end{matrix}$ е крайно, $\begin{matrix} +¿ = +\infty \\ t^{¿} \end{matrix}$.

$$-¿ \coloneqq max\left\{ \frac{-x_j}{z_j} \middle| z_j > 0 \right\} = \frac{-x_{j_0}}{z_{j_0}}, j_0 \epsilon \{1,2,..,p_0\}$$
$$t^{¿}$$

Понеже z има само + координати и нулеви; при това Az=0, то z е посока. $\begin{matrix} -¿ z \\ x' = x + t^{¿} \end{matrix}$ има по- малко от p координати и следователно има представане

$x' = \sum_{i\epsilon I} \lambda_i' x_i + d'$. Тогава $x = \sum_{i\epsilon I} \lambda_i' x_i + d' - t^{¿}$ $-¿ z$. Ясно е, че $\begin{matrix} -¿ z \\ d' - t^{¿} \end{matrix}$ е посока.

Следователно x има нужното представяне.

3. Аналогично на 2. , тук –z е посока.

2.

Теорема 1.

Нека разглеждаме (K) с КММ М. Ако М е непразно, то М има поне един връх.

Доказателство:

Да допуснем противното, а именно че множеството V от върхове е празно. Тогава по теоремата за характеризация на КММ следва, че

$\forall x \epsilon M \colon x = d_x ,$

Където d_x е или посока, или нулевия вектор. Тъй като нулевият вектор сам по себе си е връх, то остава елементите на М да са посоки. Да вземем един елемент d_x. Имаме по теоремата за характеризация на посоки, че А $d_x = 0$ и $d_x \geq 0$. Но от тук следва че в (K) вектора b е нулевия вектор и М = { x | $Ax = 0$ и $x \geq 0$ } Но тогава М съдържа нулевия вектор – противоречие.

Теорема 2.

Нека разглеждаме (K) с КММ М, което е непразно. Тогава или z(x) намалява неограничено върху М, или има връх на М, в който се достига минимум на z.

Доказателство:

Ясно е, че

1) или съществува посока d, такава, че $z(d)=c^T.d <0$
2) или за всяка посока d в М $c^T.d \geq 0$

Нека сме в случай 1). Имаме поне едно x в М. Избираме едно от посоките със свойството и да разгледаме лъча $\{ x + td \vee t \geq 0 \}$. Той принадлежи целия в М, защото d е посока. Имаме, че $z(x+td) = c^T (x+td) = c^T x + t c^T d \geq c^T x$, т.е. когато t клони към безкрайност имаме неограничено намаляване

Нека сме в случай 2). За всяко x от М имаме представяне $x = \sum_{i \epsilon I} \lambda_i x_i + d$. Тогава

$$z(x) = c^T x = c^T \left(\sum_{i \epsilon I} \lambda_i x_i + d \right) = ¿$$

$$¿ \sum_{i \epsilon I} \lambda_i c^T . x_i + c^T . d \geq \sum_{i \epsilon I} \lambda_i c^T . x_i \geq \sum_{i \epsilon I} \lambda_i \min_{i \epsilon I} (c^T . x_i) = \min_{i \epsilon I} (c^T . x_i)$$

Ако означим с x_{i_0} един от върховете, в които се достига $\min_{i \in I}(c^T . x_i)$, то имаме, че за произволно x от M е в сила неравенството $z(x_{i_0}) < z(x)$, което трябваше да докажем.

Забележка. Нещата в червено са незадължителни (Златева).

Литература:

[1] Записки по МО1, спец. ПМ, Н. Златева.

Темата е разработена от Велико Дончев, уч. 2011/2012 г.