Časopis pro pěstování matematiky

Jaromír Abrham
Přibližná metoda pro nelineární programování

Časopis pro pěstování matematiky, Vol. 83 (1958), No. 4, 425--439

Persistent URL: http://dml.cz/dmlcz/108637

Časopis pro pěstování matematiky, roč. 83 (1958), Praha

PŘIBLIŽNÁ METODA PRO NELINEÁRNÍ PROGRAMOVÁNÍ

JAROMÍR ABRHAM, Praha

(Došlo dne 27. srpna 1957) DT: 512.25

V práci je konstruován konvergentní iterační postup pro určení minima ryze konvexní funkce na množině všech nezáporných řešení daného systému lineárních rovnic.

1. Úvod

Úlohou lineárního programování je najít extremní hodnoty lineární formy $f(X) = \sum_{i=1}^{n} c_i x_i$ na množině všech nezáporných řešení daného systému lineárních rovnic $\sum_{j=1}^{n} a_{ij} x_j = b_i$, $i = 1, \ldots, m < n$. Známou simplexovou metodu pro řešení této úlohy udal G. B. Dantzig — viz např. [1].

První přirozené zobecnění uvedené úlohy dostaneme, upustíme-li od požadavku, aby funkce, jejíž extrém hledáme, byla lineární. W. Prager [2] došel ke speciální úloze tohoto typu v souvislosti s penalisováním přetížení trati v tzv. dopravním problému. Dopravním problémem v lineárním programování rozumíme úlohu najít minimum lineární formy $\sum_{i=1}^{m} \sum_{j=1}^{n} c_{ij} x_{ij}$ na množině všech nezáporných matic (x_{ij}) typu $m \times n$ takových, že $\sum_{j=1}^{n} x_{ij} = a_i$, $i = 1, \ldots, m$, $\sum_{i=1}^{m} x_{ij} = b_j$, $j = 1, \ldots, n$, kde a_i, b_j jsou danákladná čísla, pro něž $\sum_{i=1}^{m} a_i = \sum_{i=1}^{n} b_j$. Prager pokládá koeficienty c_{ij} za lineární funkce proměnných x_{ij}, tj. klade $c_{ij} = d_{ij} + k_{ij} x_{ij}$, $k_{ij} > 0$. Určité zobecnění této úlohy je předmětem odst. 4 této práce.

Úkolem tohoto článku je sestrojit konvergentní iterační postup pro určení nezáporných čísel, která vyhovují danému systému lineárních rovnic a pro něž daná ryze konvexní funkce nabývá na množině všech nezáporných řešení tohoto systému svého minima. Nejprve zavedeme některá označení.

Velkými kursivními písmeny budeme značit body $(m + k)$-rozměrného euklidovského prostoru $\mathscr{E}_{m+k}$, jejichž souřadnice jsou nezáporné (to symbolicky zapíšeme ve tvaru $X \geqq \mathbf{0}$) a vyhovují pevně danému systému lineárních rovnic

$$\sum_{j=1}^{m+k} a_{ij}x_j = b_i, \quad i = 1, \ldots, m \quad (k \geqq 1)\,, \tag{1}$$

kde a_{ij}, b_i jsou pevně daná reálná čísla. Množinu všech nezáporných řešení soustavy (1) označíme $\mathfrak{M}$. Dále budeme předpokládat, že soustava (1) má aspoň jedno řešení $x_j > 0$, $j = 1, \ldots, m + k$; tento předpoklad je ekvivalentní s předpokladem, že dimense množiny $\mathfrak{M}$ je rovna k. Pro jednoduchost budeme dále vylučovat případ tzv. degenerace, tj. budeme požadovat, aby každý determinant m-tého řádu matice

$$\left\| \begin{matrix} a_{11}, \ldots\ldots a_{1,m+k}, & b_1 \\ \vdots & \vdots \\ a_{m1}, \ldots\ldots a_{m,m+k}, & b_m \end{matrix} \right\|,$$

který obsahuje prvky $b_1, \ldots, b_m$, byl různý od nuly.

Vektorem množiny $\mathfrak{M}$ nazveme libovolné řešení soustavy lineárních homogenních rovnic

$$\sum_{j=1}^{m+k} a_{ij}u_j = 0\,, \quad i = 1, \ldots, m\,. \tag{2}$$

Jak známo, vytváří všechna řešení soustavy (2) k-rozměrný vektorový prostor; obecné řešení soustavy (1) je pak dáno součtem nějakého daného řešení soustavy (1) a obecného řešení soustavy (2). Vektory množiny $\mathfrak{M}$ budeme značit malými polotučnými kursivními písmeny z konce abecedy, případně s indexy. Ve stejném smyslu jako pro body budeme i pro vektory psát $\boldsymbol{v} \geqq \mathbf{0}$. Polotučnými kursivními písmeny z počátku abecedy budeme pak značit pevně dané $m + k$-tice reálných čísel.

Je-li $\mathfrak{R}$ nějaká podmnožina množiny $\mathfrak{M}$ dimense $r (r \leqq k)$, bude symbol $\mathscr{I}(\mathfrak{R})$ značit množinu všech vnitřních bodů množiny $\mathfrak{R}$ vzhledem k metrice prostoru $\mathscr{E}_r$ vytvořeného množinou $\mathfrak{R}$.

Funkce $f(X)$ definovaná na množině $\mathfrak{M}$ je podle definice konvexní na $\mathfrak{M}$, jestliže pro libovolné číslo α takové, že $0 \leqq \alpha \leqq 1$ a pro libovolné dva body $X, Y \in \mathfrak{M}$, $X \neq Y$ platí $f(\alpha X + (1 - \alpha) Y) \leqq \alpha f(X) + (1 - \alpha) f(Y)$; jestliže při $0 < \alpha < 1$ platí vždy znaménko nerovnosti, je $f(X)$ ryze konvexní na množině $\mathfrak{M}$.

Množina $\mathfrak{M}$ je uzavřená a konvexní (viz např. [3]), nemusí však být omezená. O její omezenosti dovoluje rozhodnout

Věta 1. *Nutná a postačující podmínka pro to, aby množina $\mathfrak{M}$ nebyla omezená je, aby existoval vektor $\boldsymbol{v} = (v_1, \ldots, v_{m+k})$ množiny $\mathfrak{M}$ tak, že $v_i \geqq 0$, $i = 1, \ldots$ $\ldots, m + k$, $\sum_{i=1}^{m+k} v_i > 0$.*

Důkaz. Postačitelnost podmínky: Za předpokladů věty je pro každé $X \in \mathfrak{M}$ a pro každé $t \geqq 0$ $X + t\mathbf{v} \in \mathfrak{M}$, množina $\mathfrak{M}$ tedy není omezená.

Nutnost podmínky: Označme $\varrho(X, Y)$ vzdálenost bodů X, Y. Nechť množina $\mathfrak{M}$ není omezená. Pak existuje posloupnost $X_1, X_2, \ldots$ bodů množiny $\mathfrak{M}$ tak, že $\lim\limits_{r\to\infty} \varrho(X_r, X_1) = +\infty$. Definujme nyní vektory $\mathbf{v}_1, \mathbf{v}_2, \ldots$ množiny $\mathfrak{M}$ vztahy $\mathbf{v}_r = \dfrac{X_r - X_1}{\varrho(X_r, X_1)}$, $r = 1, 2, \ldots$ Pak $\|\mathbf{v}_r\| = \sqrt{\sum\limits_{i=1}^{m+k} v_{ri}^2} = 1$ a existuje tedy vybraná konvergentní posloupnost $\mathbf{v}_{r_1}, \mathbf{v}_{r_2}, \ldots$; označme $\mathbf{v} = \lim\limits_{s\to\infty} \mathbf{v}_{r_s}$. Pro komponenty $v_1, \ldots, v_{m+k}$ vektoru $\mathbf{v}$ pak zřejmě platí $v_i \geqq 0$, $i = 1, \ldots$ $\ldots, m + k$, $\sum\limits_{i=1}^{m+k} v_i > 0$.

Je-li množina $\mathfrak{M}$ omezená, je existence minima funkce $f(X)$ na množině $\mathfrak{M}$ zaručena. V opačném případě platí

Věta 2. *Nechť pro libovolný vektor* $\mathbf{v} \geqq 0$, $\mathbf{v} \neq 0$ *množiny* $\mathfrak{M}$ *je* $\lim\limits_{t\to\infty} f(X + t\mathbf{v}) = +\infty$ *pro každé* $X \in \mathfrak{M}$. *Pak existuje bod* $X_0 \in \mathfrak{M}$ *tak, že* $f(X_0) = \min\limits_{X \in \mathfrak{M}} f(X)$.

Důkaz. Budiž X_1 pevný bod množiny $\mathfrak{M}$. Označme $\mathfrak{N}$ množinu všech $X \in \mathfrak{M}$, pro něž $f(X) \leqq f(X_1)$. Množina $\mathfrak{N}$ je zřejmě konvexní a uzavřená; dokážeme, že je také ohraničená. Kdyby tomu tak nebylo, existovala by posloupnost bodů $X_1, X_2, \ldots$ z množiny $\mathfrak{N}$ tak, že $\lim\limits_{r\to\infty} \varrho(X_1, X_r) = +\infty$. Podobně jako při důkazu věty 1 zkonstruujme vektor $\mathbf{v} \geqq 0$, $\mathbf{v} = \lim\limits_{s\to 0} \dfrac{X_{r_s} - X_1}{\varrho(X_{r_s}, X_1)}$.
Dokážeme nyní, že $X_1 + t\mathbf{v} \in \mathfrak{N}$ pro libovolné $t > 0$. K danému $t > 0$ existuje zřejmě $s(t)$ tak, že pro $s > s(t)$ je $\alpha_s = \dfrac{t}{\varrho(X_{r_s}, X_1)} < 1$. Mysleme si nyní t pevné a omezme se jen na $s > s(t)$. Je pak

$$f(X_1 + t\mathbf{v}) = \lim_{s\to\infty} f(X_1 + t\mathbf{v}_{r_s}) = \lim_{s\to\infty} f(X_1(1 - \alpha_s) +$$
$$+ X_{r_s}\alpha_s) \leqq \lim_{s\to\infty} [(1 - \alpha_s) f(X_1) + \alpha_s f(X_{r_s})] = f(X_1)\,,$$

neboť $X_{r_s} \in \mathfrak{N} \Rightarrow f(X_{r_s}) \leqq f(X_1)$ a $\lim\limits_{s\to\infty} \alpha_s = 0$. Pro každé $t > 0$ je tedy $f(X_1 + t\mathbf{v}) \leqq f(X_1)$, což odporuje předpokladům věty.

Jelikož $\mathfrak{N}$ je uzavřená a ohraničená, existuje $X_0 \in \mathfrak{N}$ tak, že $f(X_0) = \min\limits_{X \in \mathfrak{N}} f(X)$. Pro $X \in \mathfrak{M} - \mathfrak{N}$ je $f(X) > f(X_1) \geqq f(X_0)$, tedy $f(X_0) = \min\limits_{X \in \mathfrak{M}} f(X)$. Tím je věta dokázána.

Je-li funkce $f(X)$ ryze konvexní, což budeme většinou předpokládat, je bod, v němž nabývá na množině $\mathfrak{M}$ svého minima, určen jednoznačně. Předpokládáme-li totiž, že funkce $f(X)$ nabývá na množině $\mathfrak{M}$ svého minima ve dvou různých bodech X_0, Y_0, dostaneme

$$f(\tfrac{1}{2}X_0 + \tfrac{1}{2}Y_0) < \tfrac{1}{2}f(X_0) + \tfrac{1}{2}f(Y_0) = f(X_0)\,,$$

což je ve sporu s předpokladem $f(X_0) = \min\limits_{X \in \mathfrak{M}} f(X)$.

V dalším budeme hodně pracovat s vektory množiny $\mathfrak{M}$. Poněvadž existuje nekonečně mnoho takových vektorů i nekonečně mnoho basí vektorového prostoru popsaného rovnicemi (2), budeme dále s výjimkou odst. 4 uvažovat pouze ty vektory, které mají alespoň $k - 1$ nulových komponent. Takové vektory budeme dále nazývat základní. Pak až na číselný násobek existuje jen konečně mnoho basí uvažovaného vektorového prostoru, složených ze základních vektorů. V dalším, pokud nebude výslovně řečeno nic jiného, myslíme vektorem jen základní vektor a basí jen basi složenou ze základních vektorů.

Poznámka. Geometricky můžeme množinu $\mathfrak{M}$ chápat jako k-rozměrný uzavřený konvexní polyedr v $m + k$-rozměrném euklidovském prostoru, který obecně nemusí být omezen. Body, které mají jen kladné souřadnice, jsou pak zřejmě vnitřními body tohoto polyedru vzhledem k prostoru $\mathscr{E}_k$, popsanému systémem rovnic (1). Odtud plyne, že každý bod, který leží na hranici polyedru $\mathfrak{M}$, má alespoň jednu souřadnici rovnou nule.

2. Kriterium pro minimum

Definice 1. *Budiž $X \in \mathfrak{M}$ a $\mathbf{v}$ libovolný (ne nutně základní) vektor množiny $\mathfrak{M}$. Označme $\mathfrak{M}(\mathbf{v}, X)$ množinu všech reálných čísel t, pro něž $X + t\mathbf{v} \in \mathfrak{M}$. Pak řekneme, že bod X je minimální vzhledem k vektoru $\mathbf{v}$, jestliže $f(X) = \min\limits_{t \in \mathfrak{M}(v, X)} f(X + t\mathbf{v})$.*

Definice 2. *r-rozměrnou stěnou množiny $\mathfrak{M}$ ($0 \leqq r \leqq k$) rozumíme množinu všech bodů $X = (x_1, \ldots, x_{m+k}) \in \mathfrak{M}$, pro něž při pevně daných indexech $i_1, \ldots, i_{k-r}$* platí *$x_{i_j} = 0$, $j = 1, \ldots, k - r$, $x_s > 0$, $s \neq i_j$, $j = 1, \ldots, k - r$.*

Definice 3. *Budiž $\mathfrak{M}(r)$ ($0 \leqq r \leqq k$) r-rozměrná stěna množiny $\mathfrak{M}$, pro jejíž body je $x_{i_j} = 0$, $j = 1, \ldots, k - r$. Budiž X libovolný vnitřní bod této stěny. Pak basí příslušnou k bodu X rozumíme množinu k lineárně nezávislých vektorů $\mathbf{v}_1, \ldots, \mathbf{v}_k (\mathbf{v}_i = (v_{i,1}, \ldots, v_{i,m+k}))$ těchto vlastností:*

1. Mezi vektory $\mathbf{v}_1, \ldots, \mathbf{v}_k$ existuje právě r vektorů $\mathbf{v}_{\nu_1}, \ldots, \mathbf{v}_{\nu_r}$ tak, že $v_{\nu_s, i_j} = 0$, $s = 1, \ldots, r$, $j = 1, \ldots, k - r$

2. jsou-li $\mathbf{v}_{\mu_1}, \ldots, \mathbf{v}_{\mu_{k-r}}$ ostatní vektory této base, pak ke každému z čísel i_j, $j = 1, \ldots, k - r$ existuje právě jedno číslo μ_j, $j = 1, \ldots, k - r$ tak, že $v_{\mu_j, i_j} > 0$, $j = 1, \ldots, k - r$ a že $v_{\mu_j, i_s} = 0$, kdykoli $j \neq s$, $s = 1, \ldots, k - r$.

Věta 3. *Budiž $f(X)$ konvexní funkce definovaná na množině $\mathfrak{M}$, která má na* *množině $\mathfrak{M}$ spojité parciální derivace nejméně druhého řádu vůči všem proměnným. Pak nutná a postačující podmínka k tomu, aby bylo $f(X_0) = \min_{X \in \mathfrak{M}} f(X)$ je, aby bod X_0 byl minimální vůči všem vektorům některé k němu příslušné base.* (Přitom ty vektory base, která splňují podmínku 1 v definici 3 nemusí být základní.)

Důkaz. Nutnost podmínky je zřejmá.

Postačitelnost podmínky: Nechť X_0 je vnitřním bodem r-rozměrné stěny $\mathfrak{M}(r)$ množiny $\mathfrak{M}$, pro jejíž body platí $x_{i_1} = \ldots = x_{i_{k-r}} = 0$. (Pro $r = k$ klademe $\mathfrak{M}(k) = \mathscr{I}(\mathfrak{M})$.) Buďtež $\mathbf{v}_1, \ldots, \mathbf{v}_k$ vektory base příslušné k bodu X_0. Bez újmy obecnosti lze předpokládat, že vektory $\mathbf{v}_1, \ldots, \mathbf{v}_r$ splňují podmínku 1. definice 3 a vektory $\mathbf{v}_{r+1}, \ldots, \mathbf{v}_k$ podmínku 2. této definice. (Vektory $\mathbf{v}_1, \ldots, \mathbf{v}_r$ nemusí tedy být základní.) Budiž X libovolný bod množiny $\mathfrak{M}$. Pak lze jednoznačně psát $X = X_0 + \sum_{i=1}^{k} \gamma_i \mathbf{v}_i$, kde (vzhledem k podmínce $X \geqq 0$) $\gamma_i \geqq 0$, $i = r + 1, \ldots, k$. Položme $f(X_0 + \sum_{i=1}^{k} \gamma_i \mathbf{v}_i) = \psi(\gamma_1, \ldots, \gamma_k)$. Snadno zjistíme, že funkce $\psi(\gamma_1, \ldots, \gamma_k)$ má pak spojité parciální derivace nejméně druhého řádu na množině $\mathfrak{N}$ všech k-tic reálných čísel $(\gamma_1, \ldots, \gamma_k)$ takových, že $X_0 + \sum_{i=1}^{k} \gamma_i \mathbf{v}_i \in \mathfrak{M}$ a že je na této množině konvexní. Použitím Taylorova rozvoje dostaneme

$$\psi(\gamma_1, \ldots, \gamma_k) = \psi(0, \ldots, 0) + \sum_{i=1}^{k} \gamma_i \frac{\partial}{\partial \gamma_i} \psi(0, \ldots, 0) + R_2 ;$$

jelikož funkce $\psi(\gamma_1, \ldots, \gamma_k)$ je konvexní, je $R_2 \geqq 0$ (viz [4], str. 80). Označíme-li ještě $\varphi_i(t) = f(X_0 + t\mathbf{v}_i)$, je zřejmě $\varphi_i(t) = \psi(0, \ldots, t, \ldots, 0)$. Pro $i = r + 1, \ldots$ $\ldots, k$ máme

$$\frac{\partial}{\partial \gamma_i} \psi(0, \ldots, 0) = \frac{\mathrm{d}}{\mathrm{d}t} \varphi_i(t)\Big|_{t=0} \geqq 0 ,$$

neboť bod X_0 je minimální vůči vektoru $\mathbf{v}_i$, což znamená, že funkce $\varphi_i(t)$ je rostoucí pro každé $t > 0$, pro něž $X + t\mathbf{v}_i \in \mathfrak{M}$. (Připomínáme, že pro $t < 0$ a $r < i \leqq k$ $X_0 + t\mathbf{v}_i$ non $\in \mathfrak{M}$.) Pro ostatní i je $\frac{\partial}{\partial \gamma_i} \psi(0, \ldots, 0) = \frac{\mathrm{d}}{\mathrm{d}t} \varphi_i(t)\Big|_{t=0} = 0$; pro tato i je totiž pro všechna v absolutní hodnotě dosti malá t $X_0 + t\mathbf{v}_i \in \mathfrak{M}$ a bod X_0 je tedy minimální vůči vektoru $\mathbf{v}_i$, má-li funkce $\varphi_i(t)$ pro $t = 0$ lokální minimum; za našich předpokladů je k tomu nutnou a postačující podmínkou právě rovnost $\frac{\mathrm{d}}{\mathrm{d}t} \varphi_i(t)\Big|_{t=0} = 0$.

Odtud již plyne $\psi(\gamma_1, \ldots, \gamma_k) \geqq \psi(0, \ldots, 0)$, j. b. d.

3. Iterační postup pro určení minima ryze konvexní funkce $f(X)$ na množině $\mathfrak{M}$

V celém tomto odstavci budeme předpokládat, že je pevně dána ryze konvexní funkce $f(X)$, jejíž minimum na množině $\mathfrak{M}$ všech nezáporných řešení soustavy (1) chceme určit. Podáme nyní formulaci iteračního postupu pro určení bodu, v němž funkce $f(X)$ nabývá na množině $\mathfrak{M}$ svého minima. Budeme při tom předpokládat, že ke všem vnitřním bodům jedné a téže stěny množiny $\mathfrak{M}$ přísluší táž, pevně zvolená base, a že pořadí vektorů v každé basi je pevně určeno. Rovněž délku každého základního vektoru pokládáme za pevnou (tj. vylučujeme násobení vektorů čísly během výpočtu).

Budiž X_1 libovolný bod množiny $\mathfrak{M}$, $\mathbf{u}_1, \ldots, \mathbf{u}_k$ buďtež vektory base příslušné k bodu X_1. Položme $X_2 = X_1 + t_1\mathbf{u}_1$, kde t_1 je jednoznačně určeno podmínkou $f(X_1 + t_1\mathbf{u}_1) = \min_{t \in \mathfrak{M}(u_1, X_1)} f(X_1 + t\mathbf{u}_1)$. Přísluší-li k bodu X_2 táž base jako k bodu X_1, položme $X_3 = X_2 + t_2\mathbf{u}_2$, kde t_2 je určeno podmínkou $f(X_2 + t_2\mathbf{u}_2) = \min_{t \in \mathfrak{M}(u_2, X_2)} f(X_2 + t\mathbf{u}_2)$. Nepřísluší-li k bodu X_2 táž base jako k bodu X_1, opakujeme pro X_2 postup jako pro X_1 s prvním vektorem base příslušné k bodu X_2.

Předpokládejme, že již byl konstruován bod X_n. Buďtež $\mathbf{v}_1^{(n-1)}, \ldots, \mathbf{v}_k^{(n-1)}$ vektory base příslušné k bodu X_{n-1} a nechť $X_n = X_{n-1} + t_{n-1}\mathbf{v}_r^{(n-1)}(1 \leqq r \leqq k)$; přísluší-li k bodu X_n táž base jako k bodu X_{n-1}, položíme $X_{n+1} = X_n + t_n\mathbf{v}_{\bar{r}}^{(n-1)}$, kde $\bar{r} \equiv r + 1 \pmod{k}$, $1 \leqq \bar{r} \leqq k$ a t_n je určeno podmínkou

$$f(X_n + t_n\mathbf{v}_{\bar{r}}^{(n-1)}) = \min_{t \in \mathfrak{M}(v_{\bar{r}}^{(n-1)}, X_n)} f(X_n + t\mathbf{v}_{\bar{r}}^{(n-1)}) .$$

Nepřísluší-li k bodu X_n táž base jako k bodu X_{n-1}, opakujeme pro X_n celý postup s prvým vektorem příslušné base podobně jako u bodu X_1.

Poznámka. Je-li X vnitřní bod r-rozměrné stěny $\mathfrak{M}(r)$ množiny $\mathfrak{M}$, pak base příslušná k bodu X obsahuje, jak plyne z definice 3, právě r základních vektorů rovnoběžných se stěnou $\mathfrak{M}$ a $k - r$ základních vektorů, které nejsou s touto stěnou rovnoběžné. Umístíme-li nějaký vektor z těchto $k - r$ vektorů do bodu X a zvolíme-li jeho délku dosti malou (případným znásobením dosti malým kladným číslem), bude jeho koncový bod opět ležet v polyedru $\mathfrak{M}$. Popsaný postup můžeme interpretovat geometricky asi takto:

Zvolíme v polyedru $\mathfrak{M}$ libovolně bod X_1. Pokládáme-li jednobodové množiny za otevřené, je X_1 vždy vnitřním bodem některé stěny $\mathfrak{M}(r)$ množiny $\mathfrak{M}$. (Je-li $r = 0$, je bod X_1 vrcholem polyedru $\mathfrak{M}$, je-li $r = k$, je jeho vnitřním bodem.) Budiž $\mathbf{v}_1, \ldots, \mathbf{v}_k$ base příslušná k bodu X_1. Zvolíme nyní bod X_1 za počátek vektoru $\mathbf{v}_1$. Je-li vektor $\mathbf{v}_1$ rovnoběžný se stěnou $\mathfrak{M}(r)$, jdeme od bodu X_1 po přímce s parametrickým popisem $X = X_1 + t\mathbf{v}_1$ tak dlouho, dokud na této přímce funkce $f(X)$ klesá nebo dokud nepřijdeme na hranici stěny $\mathfrak{M}(r)$. Tímto způsobem najdeme bod X_2. Je-li bod X_1 minimální vůči vektoru $\mathbf{v}_1$, což v tomto případě znamená, že je $\frac{\mathrm{d}}{\mathrm{d}t} f(X_1 + t\mathbf{v}_1) \big|_{t=0} = 0$, klademe ovšem $X_2 = X_1$. Je-li X_2 vnitřní bod stěny $\mathfrak{M}(r)$, opakujeme pro něj celý postup s vektorem $\mathbf{v}_2$; v opačném případě je nutno měnit basi.

Není-li vektor $\mathbf{v}_1$ rovnoběžný se stěnou $\mathfrak{M}(r)$, je celý postup analogický; jdeme ovšem jen ve směru vektoru $\mathbf{v}_1$, nikoli ve směru opačném. Zastavíme-li se uvnitř některé stěny vyšší dimense, lze zřejmě zachovat basi, v opačném případě je nutno zvolit basi novou.

Pomocná věta. *Pro výše popsaný iterační postup platí* $\lim_{n\to\infty} t_n = 0$.

Důkaz. Protože je jen konečný počet základních vektorů množiny $\mathfrak{M}$, stačí tvrzení věty dokázat jen pro každý takový vektor $\mathbf{v}$, pro který pro nekonečně mnoho indexů n_r, $r = 1, 2, \ldots$ platí $X_{n_r+1} = X_{n_r} + t_{n_r}\mathbf{v}$.

Posloupnost $\{f(X_n)\}_{n=1}^{\infty}$ je zřejmě nerostoucí a zdola ohraničená, existuje tedy reálné číslo c tak, že $c = \lim_{n\to\infty} f(X_n)$. Snadno se zjistí, že posloupnost $\{X_n\}_{n=1}^{\infty}$ je omezená (ve smyslu euklidovské metriky). Odtud plyne, že je omezená i posloupnost $\{t_{n_r}\}_{r=1}^{\infty}$. Předpokládejme, že $\lim_{r\to\infty} t_{n_r} \neq 0$. Pak existuje vybraná posloupnost $\{t_{n_{r_s}}\}_{s=1}^{\infty}$ tak, že $\lim_{s\to\infty} t_{n_{r_s}} = t \neq 0$. Posloupnost $\{X_{n_{r_s}}\}_{s=1}^{\infty}$ je ohraničená, obsahuje tedy konvergentní vybranou posloupnost $\{X_{n_{r_{s_\nu}}}\}_{\nu=1}^{\infty}$. Pišme pro stručnost $Y_\nu = X_{n_{r_{s_\nu}}}$, $Z_\nu = X_{n_{r_{s_\nu}}+1}$, $\tau_\nu = t_{n_{r_{s_\nu}}}$, $\nu = 1, 2, \ldots$ Označme $Y_0 = \lim_{\nu\to\infty} Y_\nu$. Zřejmě existuje pak $Z_0 = \lim_{\nu\to\infty} Z_\nu = Y_0 + t\mathbf{v} \neq Y_0$ a platí

$$f(Y_0) = f(Z_0) = c\,, \quad f(\tfrac{1}{2}Y_0 + \tfrac{1}{2}Z_0) < \tfrac{1}{2}f(Y_0) + \tfrac{1}{2}f(Z_0) = c\,.$$

Na druhé straně je však zřejmě pro všechna ν

$$f(Z_\nu) \leqq f(\tfrac{1}{2}Y_\nu + \tfrac{1}{2}Z_\nu) \leqq f(Y_\nu)\,,$$

tedy

$$f(\tfrac{1}{2}Y_0 + \tfrac{1}{2}Z_0) = \lim_{\nu\to\infty} f(\tfrac{1}{2}Y_\nu + \tfrac{1}{2}Z_\nu) = c\,,$$

což dává spor.

Věta 4. *Posloupnost* $\{X_n\}_{n=1}^{\infty}$, *konstruovaná na začátku tohoto odstavce je konvergentní. Označíme-li* $X_0 = \lim_{n\to\infty} X_n$, *je* $f(X_0) = \min_{X\in\mathfrak{M}} f(X)$.

Důkaz. Posloupnost $\{X_n\}_{n=1}^{\infty}$ je ohraničená a má tedy alespoň jeden hromadný bod X_0. Rozlišíme nyní několik zvláštních případů.

1. Nechť $X_0 \in \mathscr{I}(\mathfrak{M})$. Budiž $\{X_{n_s}\}_{s=1}^{\infty}$ taková vybraná posloupnost posloupnosti $\{X_n\}_{n=1}^{\infty}$, že $\lim_{s\to\infty} X_{n_s} = X_0$. Protože je jen konečně mnoho základních vektorů, existuje vektor $\mathbf{w}_1$ z base příslušné ke všem bodům X_{n_s} tak, že pro nekonečně mnoho s, např. pro $s = s_\nu$, $\nu = 1, 2, \ldots$ je $X_{n_{s_\nu}+1} = X_{n_{s_\nu}} + t_{n_{s_\nu}}\mathbf{w}_1$. Označme pro stručnost $X_{n_{s_\nu}+p} = Z_\nu^{(p)}$, $t_{n_{s_\nu}+p} = \tau_\nu^{(p)}$, $p = 0, 1, \ldots, k$, $\nu = 1, 2, \ldots$ Z podmínky $\lim_{\nu\to\infty} \tau_\nu^{(1)} = 0$ plyne $\lim_{\nu\to\infty} Z_\nu^{(1)} = X_0$. Budiž $\mathbf{w}_2$ vektor následující za vektorem $\mathbf{w}_1$ v basi příslušné ke všem bodům $Z_\nu^{(1)}$ pro dosti velká ν. Pak

$Z_\nu^{(2)} = Z_\nu^{(1)} + \tau_\nu^{(1)} \mathbf{w}_2$ a opět tedy $\lim_{\nu\to\infty} Z_\nu^{(2)} = X_0$. Opakováním celého postupu najdeme $\lim_{\nu\to\infty} Z_\nu^{(p)} = X_0$, $p = 1, \ldots, k$. Kdyby nebylo $f(X_0) = \min_{X\in\mathfrak{M}} f(X)$, existoval by vektor $\mathbf{w}_i$ patřící k dané basi a číslo $t_i \neq 0$ tak, že by bylo

$$X_0 + t_i \mathbf{w}_i \in \mathfrak{M},\ f(X_0 + t_i \mathbf{w}_i) < f(X_0)\,.$$

Pak by ovšem bylo

$$\lim_{\nu\to\infty} f(Z_\nu^{(i)} + t_i \mathbf{w}_i) = f(X_0 + t_i \mathbf{w}_i) < f(X_0)$$

a tedy by bylo pro všechna dosti velká ν

$$f(Z_\nu^{(i)} + t_i \mathbf{w}_i) < f(Z_\nu^{(i+1)}) = f(Z_\nu^{(i)} + \tau_\nu^{(i)} \mathbf{w}_i)\,,$$

což je ve sporu s konstrukcí bodu $Z_\nu^{(i+1)}$. Z jednoznačnosti minima funkce $f(X)$ na množině $\mathfrak{M}$ plyne, že posloupnost $\{X_n\}_{n=1}^\infty$ nemůže mít na množině $\mathfrak{M}$ jiný hromadný bod. Kdyby totiž byl bod $Y_0 \neq X_0$ hromadným bodem posloupnosti $\{X_n\}_{n=1}^\infty$, nastávalo by minimum funkce $f(X)$ na množině $\mathfrak{M}$ ve dvou různých bodech, což je za našich předpokladů nemožné.

2. Nechť $X_0 \in \mathscr{I}(\mathfrak{M}(r))$, kde $\mathfrak{M}(r)$ je r-rozměrná stěna množiny $\mathfrak{M}$, pro jejíž body platí $x_{i_j} = 0$, $j = 1, \ldots, k - r$ a nechť od určitého n_0 počínaje $X_n \in \mathfrak{M}(r)$. Pak ke všem bodům X_n $(n \geqq n_0)$ přísluší zřejmě táž base a podobně jako sub 1. ověříme, že bod X_0 je minimální vůči všem vektorům uvažované base $\mathbf{w}_s = (w_{s,1}, \ldots w_{s,m+k})$, ležícím ve směru dané stěny, tj. takovým, pro něž $w_{s,\,i_j} = 0$, $j = 1, \ldots, k - r$. Budiž $\mathbf{w}_j$ libovolný vektor dané base neležící ve směru stěny $\mathfrak{M}(r)$. Pak musí pro všechna dosti velká n taková, že $X_{n+1} = X_n + t_n \mathbf{w}_j$, být

$$\frac{\mathrm{d}}{\mathrm{d}t} f(X_n + t\mathbf{w}_j)|_{t=0} \geqq 0\;;$$

podobně jako sub 1 usoudíme, že X_0 je minimální vzhledem k vektoru $\mathbf{w}_j$.

3. Nechť $X_0 \in \mathscr{I}(\mathfrak{M}(r))$, kde $\mathfrak{M}(r)$ má týž význam jako sub 2, a nechť pro nekonečně mnoho indexů n_s, $s = 1, 2, \ldots$ je $X_{n_s} \in \mathfrak{M}(r)$, X_{n_s+1} non $\in \mathfrak{M}(r)$. Pak pro všechna dosti velká s opět ke všem bodům X_{n_s} přísluší táž base. Bez újmy obecnosti lze předpokládat, že $\lim_{s\to\infty} X_{n_s} = X_0$. Pak ovšem také $\lim_{s\to\infty} X_{n_s+1} = X_0$. Budiž $\mathbf{v}$ takový vektor, že pro nekonečně mnoho s je $X_{n_s+1} = X_{n_s} + t_{n_s}\mathbf{v}$. Potom ovšem $\frac{\mathrm{d}}{\mathrm{d}t} f(X_{n_s} + t\mathbf{v})_{t=0} < 0$ pro nekonečně mnoho s a limitním přechodem z vybrané posloupnosti najdeme $\frac{\mathrm{d}}{\mathrm{d}t} f(X_0 + t\mathbf{v})|_{t=0} \leqq 0$. Z podmínky $t_{n_s} \to 0 (s \to \infty)$ plyne však, že platí rovnost. Minimalita bodu X_0 vůči vektorům ležícím ve směru stěny $\mathfrak{M}(r)$ se ověří jako sub 1. Minimalita vůči vektorům $\mathbf{v}_j$ neležícím ve směru stěny $\mathfrak{M}(r)$, pro něž pro všechna dosti velká n taková, že $X_{n+1} = X_n + t_n \mathbf{v}_j$, je $\frac{\mathrm{d}}{\mathrm{d}t} f(X_n + t\,\mathbf{v}_j)|_{t=0} \geqq 0$, se ověří jako sub 2.

4. Nechť $X_0 \in \mathscr{I}(\mathfrak{M}(r))$ a nechť pro všechna dosti velká n je X_n non $\in \mathfrak{M}(r)$. Protože množina $\mathfrak{M}$ má jen konečně mnoho stěn, existuje $r' > r$ a stěna $\mathfrak{M}(r')$ incidentní se stěnou $\mathfrak{M}(r)$ (tj. taková, že všechny nulové souřadnice jejích bodů jsou též nulovými souřadnicemi bodů stěny $\mathfrak{M}(r)$) tak, že pro nekonečně mnoho n, řekněme pro $n = n_s$, $s = 1, 2, \ldots$, je $X_{n_s} \in \mathscr{I}(\mathfrak{M}(r'))$. Snadno se zjistí, že, je-li X_0 hromadný bod posloupnosti $\{X_{n_s}\}_{s=1}^{\infty}$, je pro každý vektor $\mathbf{v}$ ležící ve směru dané stěny $\frac{d}{dt} f(X_0 + t\mathbf{v})|_{t=0} = 0$. Odtud plyne podobně jako v důkazu věty 3, že bod X_0 je minimální vůči všem vektorům ležícím ve směru stěny $\mathfrak{M}(r)$. Zároveň je ověřeno, že bod X_0 je minimální vůči vektorům, které neleží ve směru stěny $\mathfrak{M}(r)$, ale leží ve směru stěny $\mathfrak{M}(r')$. Pro vektory, neležící ve směru stěny $\mathfrak{M}(r')$ je minimalita bodu X_0 zřejmá.

4. Aplikace na kvadratické programování

Z chování posloupnosti $\{X_n\}_{n=1}^{\infty}$ můžeme velmi často intuitivně soudit, v které stěně množiny $\mathfrak{M}$ bude ležet bod $X_0 = \lim_{n\to\infty} X_n$. Je-li

$$f(X) = (\mathbf{d}, X) + X\mathbf{C}X' = \sum_{i=1}^{m+k} \mathrm{d}_i x_i + \sum_{i=1}^{m+k} \sum_{j=1}^{m+k} c_{ij} x_i x_j\,,$$

kde $X\mathbf{C}X'$ je positivně definitní kvadratická forma, můžeme těchto intuitivně získaných informací často využít k tomu, abychom výše studovaný nekonečný iterační postup převedli na iterační postup končící po nejvýše k krocích.

Pokud v dalším bude řeč o vektorech množiny $\mathfrak{M}$, nemusí to být vektory základní.

Definice 4. *Buďtež $\mathbf{v}_1, \ldots, \mathbf{v}_k$ lineárně nezávislé vektory, $\mathbf{C}$ positivně definitní matice. Pak vektory $\mathbf{v}_1, \ldots, \mathbf{v}_k$ jsou vzájemně ortogonální vzhledem k matici $\mathbf{C}$, je-li $\mathbf{v}_i\mathbf{C}\mathbf{v}_j' = 0$, kdykoli $i \neq j$.*

Věta 5. *Nechť všechny členy posloupnosti konstruované podle uvedeného iteračního postupu jsou vnitřní body téže stěny $\mathfrak{M}(r)$ množiny $\mathfrak{M}$ a nechť též $X_0 = \lim_{n\to\infty} X_n \in \mathscr{I}(\mathfrak{M}(r))$. Buďtež $\mathbf{v}_1, \ldots, \mathbf{v}_r$ lineárně nezávislé vektory ležící ve směru stěny $\mathfrak{M}(r)$, které jsou vzájemně ortogonální vzhledem k matici $\mathbf{C}$. Označme $X_0 = X_1 + \sum_{i=1}^{r} \gamma_i \mathbf{v}_i$. Pak, klademe-li $Y_1 = X_1 + t_1\mathbf{v}_1, \ldots, Y_r = X_1 + t_r\mathbf{v}_r$, kde čísla t_i, $i = 1, \ldots, r$ jsou konstruována způsobem popsaným v odst. 3, je $t_i = \gamma_i$, $i = 1, \ldots, r$.*

Důkaz. Jest

$$f(X_0) = f(X_1 + \sum_{i=1}^{r} \gamma_i \mathbf{v}_i) = f(X_1) + \sum_{i=1}^{r} \gamma_i(\mathbf{d}, \mathbf{v}_i) + 2\sum_{i=1}^{r} \gamma_i \mathbf{v}_i \mathbf{C} X_1' + \sum_{i=1}^{r} \gamma_i^2 \mathbf{v}_i \mathbf{C} \mathbf{v}_i'\,;$$

čísla γ_i jsou určena podmínkami $\frac{\partial}{\partial \gamma_i} f(X_1 + \sum_{i=1}^{r} \gamma_i \mathbf{v}_i) = 0$, $i = 1, \ldots, r$ a jsou tedy řešením rovnic

$$(\boldsymbol{d}, \boldsymbol{v}_i) + 2\boldsymbol{v}_i\boldsymbol{C}X_1' + 2\gamma_i\boldsymbol{v}_i\boldsymbol{C}\boldsymbol{v}_i' = 0,\ i = 1, \ldots, r\,;$$

dále je

$$f(Y_i) = f(X_1 + t_i\boldsymbol{v}_i) = f(X_1) + t_i\,(\boldsymbol{d}, \boldsymbol{v}_i) + 2t_i\boldsymbol{v}_i\boldsymbol{C}X_1' + t_i^2\boldsymbol{v}_i\boldsymbol{C}\boldsymbol{v}_i',\ i = 1, \ldots, r\,,$$

kde t_i je určeno podmínkou $\frac{\mathrm{d}}{\mathrm{d}t_i} f(X_i + t_i\boldsymbol{v}_i) = 0$, tj.

$$(\mathrm{d}, \boldsymbol{v}_i) + 2\boldsymbol{v}_i\boldsymbol{C}X_1' + 2t_i\boldsymbol{v}_i\boldsymbol{C}\boldsymbol{v}_i' = 0\,,$$

tedy $t_i = \gamma_i$, $i = 1, \ldots, r$.

Splnění předpokladů věty je těžko ověřitelné v praxi; lze ji tedy použít v určitém stadiu iteračního postupu bez ověření předpokladů a podle ní určený bod ověřit podle věty 3.

Postup, kterým z dané base vytvoříme basi, složenou z vektorů vzájemně ortogonálních vzhledem k matici $\boldsymbol{C}$, je zcela analogický postupu, vyloženém např. v [5], kap. II, § 17.

5. Příklad

Vyložené metody použijeme nyní k určení minima kvadratické formy $f(X) = 2x_1^2 + x_2^2 + 3x_3^2 + x_4^2 + x_5^2$ na množině $\mathfrak{M}$ všech nezáporných řešení systému lineárních rovnic

$$\begin{aligned} x_1 + x_2 - x_3 \qquad\quad + 2x_5 &= 1\,, \\ 2x_1 \qquad + x_3 + 2x_4 + x_5 &= 4\,. \end{aligned}$$

Zvolme $X_1 = (0, 1, 0, 2, 0)$; je pak $f(X_1) = 5$. Base příslušná k bodu X_1 je složena z vektorů

$$\boldsymbol{u}_1 = (1, -1, 0, -1, 0)\,,\ \boldsymbol{u}_2 = (0, 2, 2, -1, 0),\ \boldsymbol{u}_3 = (0, -4, 0, -1, 2).$$

Položme $X_2 = X_1 + t\boldsymbol{u}_1$; z podmínky $\frac{\mathrm{d}}{\mathrm{d}t} f(X_1 + t\boldsymbol{u}_1) = 0$ najdeme $t = 0{,}75$ a tedy $X_2 = (0{,}75;\ 0{,}25;\ 0;\ 1{,}25,\ 0)$, $f(X_2) = 2{,}75$. Podobně je $X_3 = X_2 + 0{,}04412\boldsymbol{u}_2 = (0{,}75;\ 0{,}33824;\ 0{,}08824;\ 1{,}20588;\ 0)$, $f(X_3) = 2{,}71691$. Dále je $X_4 = X_3 + t\boldsymbol{u}_3$; z podmínky $\frac{\mathrm{d}}{\mathrm{d}t} f(X_3 + t\boldsymbol{u}_3) = 0$ najdeme $t = 0{,}12185$; aby však souřadnice bodu X_4 byly nezáporné, musíme vzít jen $t = 0{,}08456$. Je pak $X_4 = (0{,}75;\ 0;\ 0{,}08824;\ 1{,}12132;\ 0{,}16912)$, $f(X_4) = 2{,}43344$. Nyní musíme změnit basi, protože žádný vektor base příslušné k bodu X_4 nesmí mít (vzhledem k definici 3) druhou komponentu zápornou; zde stačí zřejmě nahradit vektor $\boldsymbol{u}_1$ vektorem $\boldsymbol{u}_1' = -\boldsymbol{u}_1$. Dále máme $X_5 = X_4 + 0{,}09467\boldsymbol{u}_1' = (0{,}65533;\ 0{,}09467;\ 0{,}08824;\ 1{,}21599;\ 0{,}16912)$, $f(X_5) = 2{,}39847$, $X_6 = X_5 + 0{,}02925\,\boldsymbol{u}_2 = (0{,}65533;\ 0{,}15317;\ 0{,}14674;\ 1{,}18674;\ 0{,}16912)$, $f(X_6) = 2{,}38393$. Položme nyní $X_7 = X_6 + t\boldsymbol{u}_3$; z podmínky $\frac{\mathrm{d}}{\mathrm{d}t} f(X_6 + t_{u3}) = 0$ najdeme $t = 0{,}06958$, vzhledem k nezápornosti souřadnic bodu X_7 můžeme položit jen $t = 0{,}0382925$. Po zaokrouhlení na pět desetinných míst najdeme $X_7 = (0{,}65533;\ 0;\ 0{,}14674;\ 1{,}14845;\ 0{,}24570)$, $f(X_7) = 2{,}30282$. Okolnost, že body X_4 i X_7 leží ve stěně množiny $\mathfrak{M}$, pro jejíž body platí $x_2 = 0$, nás vede k domněnce, že bod $\overline{X}$, pro který je $f(\overline{X}) = \min_{X \in \mathfrak{M}} f(X)$, bude také ležet v této stěně. Zkusíme proto najít minimum formy $f(X)$ v této stěně způsobem popsaným ve větě 3. Za basi směru této stěny zvolíme např. vektory $\mathbf{v}_1 = (2, 0, 2, -3, 0)$, $\mathbf{v}_2 = (-4, 0, 0, 3, 2)$. Po orthogonalisaci najdeme vektory base $\mathbf{w}_1 = \mathbf{v}_1$, $\mathbf{w}_2 = (-66,\ 0{,}50,\ 12,\ 58)$. Položíme-li $Y_i = X_7 + t_i \mathbf{w}_i$, $i = 1,2$, najdeme $t_1 = -0{,}0019452$, $t_2 = 0{,}0018489$. Souřadnice bodu $X_0 = X_7 + t_1\mathbf{w}_1 + t_2\mathbf{w}_2$ jsou pak (s přesností na pět desetinných míst) $(0{,}52941;\ 0;\ 0{,}23529;\ 1{,}17647;\ 0{,}35294)$; je přitom $f(X_0) = 2{,}23529$. Ze způsobu konstrukce bodu X_0 je patrno, že X_0 je minimální vůči vektorům $\mathbf{w}_1$ a $\mathbf{w}_2$; snadno zjistíme, že bod X_0 je též minimální vůči vektoru $\mathbf{u}'_1$. Odtud dostáváme podle věty 3, že $f(X_0) = \min_{X \in \mathfrak{M}} f(X)$.

LITERATURA

[1] *G. B. Dantzig:* Maximization of linear function of variables subject to linear inequalities, Activity analysis of production and allocation, edited by T. C. Koopmans, kap. XXI, New York 1951.

[2] *W. Prager:* On the role of congestion in transportation problems, Zeitschrift für angewandte Mathematik und Mechanik *35* (1955), 264—268.

[3] *K. Reidemeister:* Topologie der Polyeder und kombinatorische Topologie der Komplexe, Leipzig 1953.

[4] *G. H. Hardy, J. E. Littlewood, G. Pólya:* Inequalities, Cambridge 1934.

[5] *Л. А. Люстерник, В. И. Соболев:* Элементы функционального анализа, Москва 1951.

Резюме

ПРИБЛИЖЕННЫЙ МЕТОД ДЛЯ НЕЛИНЕЙНОГО ПРОГРАММИРОВАНИЯ

ЯРОМИР АБРГАМ (Jaromír Abrham), Прага

(Поступило в редакцию 27/VIII 1957 г.)

В работе описан итерационный процесс для определения минимума строго выпуклой вниз функции $f(X)$ на множестве $\mathfrak{M}$ всех неотрицательных решений системы линейных уравнений

$$\sum_{j=1}^{m+k} a_{ij} x_j = b_i\,, \quad i = 1, 2, \ldots, m,\ k \geqq 1 \tag{1}$$

ранга m. При этом мы предполагаем, что $\mathfrak{M} \neq \emptyset$.

В целях простой формулировки интерационного процесса нужно ввести некоторые понятия.

Под вектором множества $\mathfrak{M}$ мы подразумеваем решение системы линейных однородных уравнений

$$\sum_{j=1}^{m+k} a_{ij}u_j = 0\ , \quad i = 1, 2, \ldots, m\ . \tag{2}$$

Будем предполагать, что для любого вектора $\mathbf{v} = (v_1, \ldots, v_{m+k})$ множества $\mathfrak{M}$ такого, что $v_i \geqq 0$, $i = 1, 2, \ldots, m + k$, $\sum_{i=1}^{m+k} v_i > 0$ будет $\lim_{t\to\infty} f(X + t\mathbf{v}) = +\infty$ для всех $X \in \mathfrak{M}$. При этом условии в п. 1 доказывается существование минимума функции $f(X)$ на множестве $\mathfrak{M}$.

Вектор $\mathbf{u} = (u_1, \ldots, u_{m+k})$ множества $\mathfrak{M}$ является основным, если существуют индексы $i_1, \ldots, i_{k-1}$ так, что $u_{i_j} = 0$, $j = 1, \ldots, k-1$.

Пусть $X = (x_1, \ldots, x_{m+k})$ — точка множества $\mathfrak{M}$. Пусть существует целое число r $(0 \leqq r \leqq k)$ и индексы $i_1, \ldots, i_{k-r}$ такие, что $x_{i_j} = 0$, $j = 1, \ldots, k-r$. Тогда базисом, соответствующим точке X, будем разуметь множество k линейно независимых основных векторов $\mathbf{v}_1, \ldots, \mathbf{v}_k$ $(\mathbf{v}_i = (v_{i1}, \ldots, v_{i,m+k}))$, $i = 1, \ldots, k)$ со следующими свойствами:

1. Среди векторов $\mathbf{v}_1, \ldots, \mathbf{v}_k$ существует в точности r векторов $\mathbf{v}_{\nu_1}, \ldots, \mathbf{v}_{\nu_r}$ таких, что $v_{\nu_s, i_j} = 0$, $s = 1, \ldots, r$, $j = 1, \ldots, k - r$.

2. Если $\mathbf{v}_{\mu_1}, \ldots, \mathbf{v}_{\mu_{k-r}}$ — остальные векторы этого базиса, то для каждого из чисел i_j, $j = 1, \ldots, k - r$ существует в точности одно число μ_j, $j = 1, \ldots, k - r$ так, что $\mathbf{v}_{\mu_j, i_j} > 0$, $j = 1, \ldots, k - r$, и что $\mathbf{v}_{\mu_j i_s} = 0$ для всех $j \neq s$.

Точку $X \in \mathfrak{M}$ назовем минимальной относительно вектора $\mathbf{v}$, если для всякого действительного числа t такого, что $X + t\mathbf{v} \in \mathfrak{M}$, имеет место неравенство $f(X + t\mathbf{v}) \geqq f(X)$.

Для доказательства сходимости итерационного процесса является важной

Теорема 1. *Пусть $f(X)$ — выпуклая вниз функция, определенная на множестве $\mathfrak{M}$ и обладающая на множестве $\mathfrak{M}$ непрерывными частными производными первого и второго порядка по всем переменным. Тогда необходимым и достаточным условием для того, чтобы для точки X_0 было $f(X_0) = \min_{X \in \mathfrak{M}} f(X)$, является требование, чтобы точка X_0 была минимальной относительно всех векторов какого-либо соответствующего ей базиса.*

Дадим теперь формулировку итерационного процесса для отыскания минимума строго выпуклой функции $f(X)$ на множестве $\mathfrak{M}$.

Пусть X_1 — произвольная точка множества $\mathfrak{M}$, пусть $\mathbf{u}_1, \ldots, \mathbf{u}_k$ — векторы соответствующего этой точке базиса. Положим $X_2 = X_1 + t_1\mathbf{u}_1$, где t_1 однозначно определяется условием

$$f(X_1 + t_1\mathbf{u}_1) = \min_{t \in \mathfrak{M}(u_1, X_1)} f(X_1 + t\mathbf{u}_1)\ ,$$

а $\mathfrak{M}(\mathbf{u}_1, X_1)$ обозначает множество всех действительных чисел t, для которых $X_1 + t\mathbf{u}_1 \in \mathfrak{M}$. Если точке X_2 соответствует тот же базис, как и точке X_1, то положим $X_3 = X_2 + t_2\mathbf{u}_2$, где t_2 определяется условием

$$f(X_2 + t_2\mathbf{u}_2) = \min_{t \in \mathfrak{M}(u_2 X_2)} f(X_2 + t\mathbf{u}_2) .$$

Если же точке X_2 не соответствует тот же базис, что точке X_1, то для X_2 повторяем процесс, примененный к X_1, с первым вектором соответствующего точке X_2 базиса.

Предположим, что мы уже построили точку X_n. Пусть $\mathbf{v}_1^{(n-1)}, \ldots, \mathbf{v}_k^{(n-1)}$ — векторы соответствующего точке X_{n-1} базиса и пусть $X = X_{n-1} + t_{n-1}\mathbf{v}_r^{(n-1)}$ $(1 \leqq r \leqq k)$; если точке X_n соответствует тот же басис, как и точке X_{n-1}, положим $X_{n+1} = X_n + t_n\mathbf{v}_{\bar{r}}^{(n-1)}$, где $\bar{r} \equiv r + 1 \pmod{k}$, $1 \leqq \bar{r} \leqq k$ и t_n определяется условием

$$f(X_n + t_n\mathbf{v}_{\bar{r}}^{(n-1)}) = \min_{t \in \mathfrak{M}(v_{\bar{r}}^{(n-1)}, X_n)} f(X_n + t\mathbf{v}_{\bar{r}}^{(n-1)}) .$$

Если же точке X_n не соответствует тот же базис, что точке X_{n-1}, то для X_n повторяем весь процесс с первым вектором соответствующего базиса аналогично тому, как это было у точки X_1. Тогда справедлива

Теорема 2. *Последовательность* $\{X_n\}_{n=1}^{\infty}$, *построенная по описанному выше способу, является сходящейся; если обозначить* $X_0 = \lim_{n\to\infty} X_n$, *то* $f(X_0) = \min_{X \in \mathfrak{M}} f(X)$.

Параграф 4 посвящается применению изложенной теории к случаю, когда $f(X)$ является суммой линейной формы и положительно-определенной квадратичной формы.

В параграфе 5 показано применение изложенного метода на примере.

Summary

AN APPROXIMATIVE METHOD FOR NONLINEAR PROGRAMMING

JAROMÍR ABRHAM, Praha

(Received August 27, 1957)

In the present paper an iterative process is formulated for finding the minimum of a strictly convex function $f(X)$ on the set $\mathfrak{M}$ of all nonnegative solutions of a system of linear equations

$$\sum_{j=1}^{m+k} a_{ij}x_j = b_i, \quad i = 1, \ldots, m, \ k \geqq 1 \tag{1}$$

of rank m. We suppose that $\mathfrak{M} \neq \emptyset$.

Before formulating our iterative process we must introduce some concepts.

By a vector of the set $\mathfrak{M}$ we mean a solution of the system of linear homogenous equations

$$\sum_{j=1}^{m+k} a_{ij}u_j = 0\,, \quad i = 1, \ldots, m\,. \tag{2}$$

We shall suppose that $\lim_{t\to\infty} f(X + t\mathbf{v}) = +\infty$ for every vector $\mathbf{v} = (v_1, \ldots, v_{m+k})$ of $\mathfrak{M}$ such that $v_i \geqq 0$, $i = 1, \ldots, m + k$, $\sum_{i=1}^{m+k} v_i > 0$ for every $X \in \mathfrak{M}$. Under this assumption the existence of the finite minimum of $f(X)$ on $\mathfrak{M}$ is proved (sec. 1).

A non-zero vector $\mathbf{u} = (u_1, \ldots, u_{m+k})$ of $\mathfrak{M}$ is said to be *basic* if there are indices $i_1, \ldots, i_{k-1}$ such that $u_{i_j} = 0$, $j = 1, \ldots, k - 1$.

Let $X = (x_1, \ldots, x_{m+k})$ be a point of $\mathfrak{M}$ for which $x_{i_j} = 0$, $j = 1, \ldots, k - r$, $0 \leqq r \leqq k$. Then a *basis corresponding to* X is a set of k linearly independent basic vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k$ ($v_i = (v_{i1}, \ldots, v_{i,m+k})$, $i = 1, \ldots, k$) such that

1. there exist r vectors $\mathbf{v}_{\nu_1}, \ldots, \mathbf{v}_{\nu_r}$ in the basis for which $v_{\nu_s, i_j} = 0$, $s = 1, \ldots, r$, $j = 1, \ldots, k - r$.

2. if $\mathbf{v}_{\mu_1}, \ldots, \mathbf{v}_{\mu_{k-r}}$ are all others vectors of this basis then to every i_j, $j = 1, \ldots, k - r$ there exists exactly one μ_j such that $v_{\mu_j, i_j} > 0$, $j = 1, \ldots, k - r$ and that $v_{\mu_j, i_s} = 0$ for $j \neq s$.

A point $X \in \mathfrak{M}$ is *minimal with respect to a vector* $\mathbf{v}$ of $\mathfrak{M}$ if for every real t such that $X + t\mathbf{v} \in \mathfrak{M}$ it is $f(X + t\mathbf{v}) \geqq f(X)$.

For the proof of the convergence we need

Theorem 3. *Let $f(X)$ be a convex function defined on $\mathfrak{M}$ and posessing continucus partial derivatives of the 1st and 2nd order with respect to all variables. Than a necessary and sufficient condition for the relation $f(X_0) = \min_{x \in \mathfrak{M}} f(X)$ is that the point X_0 be minimal with respect to all vectors of some corresponding basis.*

We are now in a position to formulate our iterative process. Let X_1 be an arbitrary point of $\mathfrak{M}$, $\mathbf{u}_1, \ldots, \mathrm{u}_k$ the vectors of a corresponding basis. Let us put $X_2 = X_1 + t_1\mathbf{u}_1$ where t_1 is uniquely determined by the condition

$$f(X_1 + t_1\mathbf{u}_1) = \min_{t \in \mathfrak{M}(\mathbf{u}_1, X_1)} f(X + t\mathbf{u}_1)$$

($\mathfrak{M}(\mathbf{u}_1, X_1)$ is the set of all real t for which $X_1 + t\mathbf{u}_1 \in \mathfrak{M}$). If the same basis corresponds also to X_2 we put $X_3 = X_2 + t_2\mathbf{u}_2$ determining t_2 by the condition

$$f(X_2 + t_2\mathbf{u}_2) = \min_{t \in \mathfrak{M}(\mathbf{u}_2, X_2)} f(X_2 + t\mathbf{u}_2)\,.$$

If this is not the case we proceed for X_2 with the first vector of a corresponding basis as for X_1.

Let us now suppose we have already constructed the X_n. Let $\mathbf{v}_1^{(n-1)}, \ldots$ $\ldots, \mathbf{v}_k^{(n-1)}$ be the basis corresponding to X_{n-1} and let $X_n = X_{n-1} + t_{n-1}\mathbf{v}_r^{(n-1)}$ $(1 \leqq r \leqq k)$; if this basis corresponds also to X_n we put $X_{n+1} = X_n + t_n\mathbf{v}_{\bar{r}}^{(n-1)}$ where $\bar{r} \equiv r + 1 \pmod{k}$, $1 \leqq \bar{r} \leqq k$ and determine t_n by the condition

$$f(X_n + t_n\mathbf{v}_{\bar{r}}^{(n-1)} (= \min_{t \in \mathfrak{M}(v_{\bar{r}}^{(n-1)}, X_{n-1})} f(X_n + t\mathbf{v}_{\bar{r}}^{(n-1)}) .$$

In the contrary case we find a basis corresponding to X_n and construct X_{n+1} from X_n as X_2 from X_1.

Theorem 4. *The sequence $\{X_n\}_{n=1}^{\infty}$ constructed by the above described method is convergent. Denoting $X_0 = \lim_{n \to \infty} X_n$ we have $f(X_0) = \min_{x \in \mathfrak{M}} f(X)$.*

In sec. 4 the case is studied where $f(X)$ is a sum of a linear form and a positive definite quadratic form.

In sec. 5 a numerical exemple is given.