

tanulmányok

49/1976



MTA Számítástechnikai és Automatizálási Kutató Intézet Budapest



MAGYAR TUDOMÁNYOS AKADÉMIA
SZÁMITÁSTECHNIKAI ÉS AUTOMATIZÁLÁSI KUTATÓ INTÉZETE

NEMLINEÁRIS PROGRAMOZÁSI FELADATOK MEGOLDÁSA
SZEKVENCIÁLIS MÓDSZEREKKEL

Irta:

DR GERENCSÉR LÁSZLÓ

Kandidátusi értekezés

Tanulmányok 49/1976.

A kiadásért felelős:

Dr Vámos Tibor

ISBN 963 311 020 3

768015 MTA KISZ Sokszorosító. I. v.: Szabó Gyula

TARTALOMJEGYZÉK

	Oldal
<u>BEVEZETÉS</u>	5
<u>I. ELMÉLETI EREDMÉNYEK ÖSSZEFOGLALÁSA</u>	11
1. A feladat megfogalmazása	11
2. Optimalitási kritériumok	14
3. A SUMT módszer és az optimalitási kritériumok összefüggése	18
4. Néhány mátrixelméleti eredmény	20
5. Konvexitási kérdések	27
<u>II. AZ EXTRAPOLÁCIÓS TECHNIKA ÉS KITERJESZTÉSE</u>	32
1. Az extrapolációs technika megalapozása	32
2. Az extrapolációs technika kiterjesztése	37
3. A Householder-triangularizáció alkalmazása	40
4. Az extrapolációs technika további büntetőfüggvények esetén	47
5. Érzékenységi vizsgálatok	53
<u>III. A MULTIPLIKÁTORMÓDSZER</u>	58
1. Elméleti összefoglalás	58
2. A multiplikátormódszer levezetése	62
3. A bővített Lagrange-függvény további vizsgálata	68
4. A multiplikátormódszer alkalmazásainak numerikus kérdései	70
<u>IV. SZEKVENCIAÁLIS MÓDSZEREK ALKALMAZÁSA A SZTOHASZTIKUS PROGRAMOZÁSBAN</u>	74
1. A sztohasztikus approximáció néhány módszere	74
2. Veszteségfüggvényes és megbízhatósági jellegű modellek	77

<u>FÜGGELÉK</u>	85
1. Egy dekompozíciós eljárás	85
2. Sztohasztikus Newton-módszer	92
<u>IRODALOMJEGYZÉK</u>	95

BEVEZETÉS

Dolgozatomban a nemlineáris programozás néhány elméleti kérdését és a büntetőfüggvények módszerével /SUMT/ illetve a multiplikátor módszerrel kapcsolatos nehézségeket és új eredményeket tárgyalok. A közölt vizsgálatok kiindulópontja Fiacco és McCormick könyve volt, amely az első összefoglaló mű a SUMT módszerről. Számos érdekes eredmény mellett az említett könyvben sok a nyitott kérdés is, elsősorban a SUMT módszer numerikus megvalósításával kapcsolatban. A SUMT módszer alkalmazása során ugyanis igen rosszul kondicionált részfeladatokat kapunk. A rosszul kondicionáltság elhárítására két módszert mutatunk be. Az egyik módszer a segédfüggvény Hesse-mátrixának egy alkalmas felbontásán alapszik, amelyet a lineáris algebrában ismert Householder-triangularizáció segítségével valósíthatunk meg. Ezt a módszert a második fejezetben ismertetjük. A második módszer a multiplikátorok módszere, amely az irodalomban is olvasható. Ezt a harmadik fejezetben tárgyaljuk, új bizonyítással. Rámutatunk az új matematikai képalkotás elméleti következményeire is.

A dolgozat célja az új eredmények közzétételén túl egy összefoglaló tárgyalás nyújtása. Főleg azokat a kérdéseket részletezzük, amelyek Fiacco és McCormick könyvében nem szerepelnek. Ezért kaptak több helyet a dolgozat első fejezetében a kvázikonvex függvényekkel kapcsolatos eredmények és néhány központi szerepű mátrixelméleti eredmény.

A SUMT módszerrel kapcsolatos vizsgálatokat nem tekinthetjük befejezettnek. A dolgozatban közölt eredmények alkalmazása speciális feladatok esetén számos új, érdekes kérdést vet fel. Mielőtt erre rátérnénk, röviden áttekintést adunk a SUMT módszer első megjelenési formáiról.

Az első feladat egy variációs probléma:

$$\min_{\int_0^T} F(x, \dot{x}, t) dt$$

/1/

$$g_i(x) = 0 \quad i = 1, \dots, m.$$

A keresett $x(t)$ görbe kezdő és végpontja rögzített. Ilyen feladatra vezet pl. egy geodetikus vonal meghatározása. Az /1/ feladat helyett tekinthetjük a következő segédfeladatot:

$$\min_{\int_0^T} (F(x, \dot{x}, t) + \epsilon^{-1} \sum_{i=1}^m g_i^2(x)) dt.$$

Ez egy feltétel nélküli variációs probléma, amelynek ϵ -től függő $x(\epsilon, t)$ megoldása tart az /1/ feladat megoldásához, ha $\epsilon \rightarrow 0$. Az $x(\epsilon, t)$ közelítő megoldás meghatározásához meg kell oldanunk az Euler-Lagrange egyenletet, ami egy másodrendű nemlineáris differenciálegyenlet.

A SUMT módszer történetében a második jelentős alkalmazás a nemlineáris programozás feladataival kapcsolatos. A feladatot a

$$\begin{aligned} & \min f(x) \\ /2/ \quad & g_i(x) \geq 0 \quad i = 1, \dots, m \\ & h_j(x) = 0 \quad j = 1, \dots, p \end{aligned}$$

alakban adjuk meg.

A /2/ feladat helyett tekinthetjük a

$$\min P(x,r) = f(x) - r \sum_{i=1}^m \ln g_i(x) + r^{-1} \sum_{j=1}^p h_j^2(x)$$

feladatot. Ez egy feltétel nélküli minimalizálással megoldható. Az $x(r)$ megoldás r -től függ és megmutatható, hogy $r \rightarrow 0$ esetén $x(r)$ tart a /2/ feladat megoldásához. A dolgozat teljes egészében a /2/ feladattal foglalkozik.

A SUMT módszer legfrisebb alkalmazási területe az irányításelmélet. Tekintsük az

$$\dot{x} = f(x,u)$$

/3/

$$\min \int_0^T F(x,u) dt$$

feladatot, ahol $x(0)$, $x(T)$ adottak. Itt x a rendszer fázisvektora, u az irányítási vektor. A /3/ feladat helyett tekintsük a

$$\min \int_0^T (F(x,u) + \varepsilon^{-1} \| \dot{x} - f(x,u) \|^2) dt$$

variációs problémát. Az $x(0)$, $x(T)$ értékek továbbra is rögzítettek. Megmutatható, hogy $\varepsilon \rightarrow 0$ esetén a variációs probléma megoldása tart a /3/ feladat megoldásához. Irányításelméleti feladatoknak ezt a megoldási módját szokás ε technikának is nevezni, a módszer megalkotója, Balakrishnan szóhasználatát követve. Sajnos Balakrishnan munkáiban a numerikus szempontok elsikkadnak. Újabb cikkekben a várható numerikus nehézségeket a multiplikátor-módszer alkalmazásával próbálják megelőzni. Figyelemreméltó vonása a /3/ feladatnak, hogy az extrapolációs technikának a dolgozatunkban bemutatott kezelési módja nem alkalmazható. Ennek oka pedig az, hogy a Householder-triangularizációt nem sikerült lineáris differenciál-operátorokra hatékonyan alkalmazni.

A jelen vizsgálatokat a következő problémák indították el: a /2/ feladatban a $P(x,r)$ függvény minimalizálása igen nehéz, mivel $\nabla_{xx}^2 P(x,r)$ az optimum közelében csaknem szinguláris. Ezért egy extrapolációs technikát dolgozunk ki, amely az $x(r)$ görbe lineáris közelítésén alapul, s az x^* optimumot az

$$x^* \approx x(r) - r \frac{dx(r)}{dr}$$

közelítés alapján számoljuk ki. A második fejezetben megmutatjuk, hogyan alkalmazható ez a közelítés az $x(r)$ görbén kívül is, másrészt hogyan számítható ki hatékonyan a $dx(r)/dr$ érintővektor.

A multiplikátormódszernél a következő probléma merül fel: a /2/ feladathoz megszerkesztjük a

$$Q(x,w,k) = f(x) + \sum_{j=1}^p w_j h_j(x) + k \sum_{j=1}^p h_j^2(x)$$

un. bővített Lagrange függvényt. Feltesszük, hogy minden feltétel egyenlőség. Adott w mellett megoldjuk a

$$\min_x Q(x,w,k)$$

feladatot, a megoldást jelölje $x(w)$ és ezután w értékét a

$$\delta w_j = 2k h_j(x(w))$$

korrekciókkal módosítjuk. Fontos kérdés már most, hogy az $x(w)$ minimumot milyen pontossággal kell meghatározni. Megmutatjuk, hogy a w és x változók egyidejű korrekciója több módon is lehetséges. Erre példa a III. fejezet (4.13) algoritmus.

A dolgozat negyedik fejezetében megmutatjuk, hogy a sztochasztikus programozás két fontos feladattípusára, a veszteségfüggvényes és a megbízhatósági jellegű sztochasztikus programozási feladatra

a SUMT módszer ill. a multiplikátormódszer sikeresen alkalmazható. Ezt úgy értjük, hogy a feladat feltételi függvényeinek Monte-Carlo módszerrel történő kiértékelését és a nemlineáris programozási eljárásokat az esetek nagy részében sikerült egy sztohasztikus approximációs eljárássá egyesíteni.

Végül a Függelékben általános esetben is vizsgáljuk bonyolult feladatok dekompozíciójának a kérdését. A probléma a következő: adott egy

$$x_{n+1} = x_n + p(x_n, Y(x_n))$$

algoritmus, amelynek a jobboldalán egy kifejezést mint egy $Y(x_n)$ függvényt kezelünk. Ez a függvény csak algoritmikus uton határozható meg, mégpedig az

$$Y_{n+1} = Y_n + r(Y_n, x)$$

algoritmus alapján. Megmutatjuk, hogy a szóbanforgó két algoritmus egyetlen konvergens algoritmussá egyesíthető elég tág feltételek mellett. Ezt a Függelék (1.6), (1.7) képletében adjuk meg.

A dolgozat célkitűzése hármasszerű. Elsőként az irodalomban hiányosan, hibásan vagy kellő elegancia nélkül közölt eredményeknek kívánok egy egységes és viszonylag rövid tárgyalást adni. Másodsor a közvetlen környezetemben, illetve az irodalomban felmerülő néhány megoldatlan számítástechnikai probléma megoldását mutatom be. Az új módszerek így a meglévő matematikai programozási software továbbfejlesztésére könnyen alkalmazhatók. A közölt eljárások alapvetően új módszertani elemeket tartalmaznak, a hangsúly is ezen van. A minden részletre kiterjedő algoritmikus sémákat is ezért mellőzöm. Végül a dolgozat célja újabb vizsgálatok kezdeményezése, amennyiben az új metodika alkalmazása az optimalizálás más területein /játékelmélet, folyamatirányítás/ jelentős erőfeszítéseket kíván.

A dolgozatnak a hazai, illetve nemzetközi kutatásokban elfoglalt helyét a főbb hivatkozások megadásával kívánom megvilágítani. Az irodalomjegyzék nem kimerítő, de úgy vélem, elegendő segítséget nyújt az érdeklődő olvasó számára.

A dolgozat megírásában nagy segítségemre voltak az MTA SZTAKI Operációkutatási Osztályának munkatársai, különösen pedig Prékopa András és a nemlineáris programozási csoport tagjai: Bernau Heinz, Mayer János, Rapcsák Tamás, valamint Kutas Tibor. Hálával tartozom a moszkvai VCAN és az oxfordi egyetem Computing Laboratory munkatársainak is, akiktől sokat tanultam. Köszönettel tartozom Baján Lászlónénak és Gabnai Katalinnak a gépeléséért és a Tudományos Titkárságnak a dolgozat kiállításában nyújtott segítségéért.

I. ELMÉLETI EREDMÉNYEK ÖSSZEFOGLALÁSA

1. A feladat megfogalmazása

A nemlineáris programozás feladatát a

$$\begin{array}{l} \text{/A/} \quad \min f(x) \\ \quad \quad g_i(x) \geq 0 \quad i = 1, \dots, m \end{array}$$

alakban fogalmazzuk meg. Itt $x \in E^n$, ahol E^n az n -dimenziós Euklidesi teret jelöli. A feltételek által meghatározott tartományt R -rel jelöljük. Feltesszük, hogy R nem üres kompakt halmaz. Az $f(x)$, $g_i(x)$ függvényekről feltesszük, hogy differenciálhatók egy az R halmazt tartalmazó nyílt halmazon. Definiáljuk az

$$R^{\circ} = \{x \mid g_i(x) > 0, \quad i = 1, \dots, m\}$$

halmazt. Megköveteljük, hogy R° lezártja R° legyen. Feltételeink mellett az $f(x)$ függvény R -en eléri a minimumát. Feltesszük, hogy az (A) feladatnak egyetlen lokális megoldása van, amelyet x^* -gal jelölünk. Ez a feltevés általában teljesül, ha a feladat egy lokális minimumára már ismert egy közelítés és a megengedett tartományt új feltételek hozzávételével szűkítjük. Feladatunk x^* meghatározása.

Az (A) feladatnál általánosabb a következő (B) feladat

$$\begin{array}{l} \text{/B/} \quad \min f(x) \\ \quad \quad g_i(x) \geq 0 \quad i = 1, \dots, m \\ \quad \quad h_j(x) = 0 \quad j = 1, \dots, p \end{array}$$

Az f, g_i, h_j függvények differenciálhatók egy az S halmazt tartalmazó nyílt halmazon. /A megengedett tartományt S -sel jelöljük./ Feltesszük, hogy S kompakt halmaz.

Az egyenlőtlenség feltételek által meghatározott halmazzt R -rel jelöljük. Az (A) feladathoz hasonlóan feltesszük, hogy R^0 lezártja R , továbbá hogy az $f(x)$ függvénynek az S halmazon egyetlen lokális minimuma van, amit x^* -gal jelölünk.

Mind az (A) mind pedig a (B) feladat esetén használni fogjuk a

$$I(x^*) = \{ i \mid g_i(x^*) = 0 \}$$

jelölést. $I(x^*)$ tehát az aktiv indexek halmaza.

Az (A) (B) feladatok megoldására elterjedt a SUMT módszer. A SUMT módszer részletes kidolgozása Fiacco és McCormick érdeme. Röviden áttekintjük a SUMT módszereket. Alapgondolatunk a következő: A feladathoz szerkesztünk egy $P(x,r)$ segédfüggvényt, amelynek feltétel nélküli minimuma $x(r)$ konvergál x^* -hoz, amint $r \rightarrow 0$. A segédfüggvényeket ugynevezett büntetőfüggvények segítségével szerkesztjük. Első példánk a logaritmikus büntetőfüggvény, amelyet az (A) feladat megoldására használunk. Definíciója:

$$/1.1/ \quad I(x) = - \sum_{i=1}^m \ln g_i(x).$$

A segédfüggvényt a

$$P(x,r) = f(x) - r \sum_{i=1}^m \ln g_i(x)$$

képlettel számítjuk. Ez a függvény R -n felveszi a minimumát, mert R határához közeledve $P(x,r)$ tart végtelenhez. Egy minimum helyet jelöljön $x(r)$. Belátható, hogy $r \rightarrow 0$ esetén $x(r)$ tart x^* -hoz. Mivel $P(x,r)$ általában nem konvex az $x(r)$ meghatározása nem triviális feladat.

Ugyancsak az (A) feladat megoldásánál használjuk a reciprok büntető függvényt, amelynek definíciója

$$/1.3/ \quad I(x) = \sum_{i=1}^m \frac{1}{g_i(x)}$$

A négyzetes büntetőfüggvény alkalmazására akkor van szükség, amikor a feltételek között egyenlőségek is szerepelnek. A négyzetes büntetőfüggvényt a (B) feladat esetén a következő módon definiáljuk:

$$/1.4/ \quad I(x) = \sum_{i=1}^m g_i(x)_-^2 + \sum_{j=1}^p h_j^2(x),$$

ahol $g_i(x)_- = \min \{0, g_i(x)\}$, a segédfüggvényt pedig a

$$/1.5/ \quad P(x, r) = f(x) + \frac{1}{r} I(x)$$

képlettel számítjuk. A négyzetes büntetőfüggvény alkalmazásának egyik előnye, hogy nem kívánja megengedett pont ismeretét. A módszert külső pont módszernek is nevezik, megmutatható ugyanis, hogy a $P(x, r)$ függvény feltétel nélküli minimuma általában R -en kívül esnek. A (B) feladat megoldásának egy másik útja vegyes büntetőfüggvények alkalmazása. Ezt a következőképpen szerkesztjük:

$$/1.6/ \quad P(x, r) = f(x) - r \sum_{i=1}^m \ln g_i(x) + \frac{1}{r} \sum_{j=1}^p h_j^2(x).$$

Vegyes módszert kaphatunk a reciprok büntetőfüggvény és a négyzetes büntetőfüggvény kombinálásából is. A segédfüggvény helyes alakja ekkor

$$/1.7/ \quad P(x, r) = f(x) + r \sum_{i=1}^m \frac{1}{g_i(x)} + r^{-\frac{1}{2}} \sum_{j=1}^p h_j^2(x).$$

Az (A) feladat esetén a négyzetes büntetőfüggvény alkalmazásának egyik hátránya az, hogy második deriváltja R határán nem létezik. Ezt a hátrányt kiküszöböli a következő exponenciális büntetőfüggvény:

$$/1.8/ \quad P(x,r) = f(x) + r \sum_{i=1}^m e^{-\frac{1}{r} g_i(x)} .$$

A konvergencia tételt az általánosabb (B) feladatra fogalmazzuk meg.

1. Tétel. A (B) feladathoz szerkesszük meg az (1.4), (1.6), (1.7) $P(x,r)$ segédfüggvények valamelyikét.

A $P(x,r)$ függvény feltételnélküli minimumát jelölje $x(r)$.

A (B) feladattal kapcsolatban mondott feltételek mellett $x(r) \rightarrow x^*$ amint $r \rightarrow 0$.

2. Optimalitási kritériumok

Először az (A) feladattal kapcsolatban fogalmazzunk meg optimalitási kritériumokat.

Megköveteljük a következő feltétel teljesülését:

$$(\alpha) \text{ az } x^* \text{ pontban a } \nabla g_i(x^*) \quad i \in J(x^*)$$

vektorok lineárisan függetlenek. Ekkor teljesül a Kuhn-Tucker féle elsőrendű regularitási feltétel [ld. [27]]. Ebből pedig következik, hogy x^* -ban teljesül az optimalitás Kuhn-Tucker szükséges feltétele, amelyet a következő tételben fogalmazzunk meg.

2. Tétel. Ha az (A) feladatra teljesül az (α) feltétel, akkor léteznek olyan u_i^* Lagrange-szorozók, hogy fennállnak a

$$/2.1/ \quad \nabla f(x^*) - \sum_{i=1}^m u_i^* \nabla g_i(x^*) = 0$$

$$u_i^* g_i(x^*) = 0 \quad i = 1, \dots, m$$

$$u_i^* \geq 0$$

feltételek.

A második feltétel elnevezése komplementaritási feltétel.

Ha itt u_i^* és $g_i(x^*)$ egyidejűleg nem nulla bármely i -re, akkor azt mondjuk, hogy a szigorú komplementaritási feltétel teljesül.

Vezessük be a Lagrange függvényt az

$$/2.2/ \quad L(x, u) = f(x) - \sum_{i=1}^m u_i g_i(x)$$

képlettel. Itt u egy m -dimenziós vektort jelöl, melynek i -dik komponense u_i . A Lagrange-függvényt csak $u_i \geq 0$ mellett értelmezzük. A /2.1/ feltételeket így is kifejezhetjük

$$/2.2/ \quad \nabla_x L(x^*, u^*) = 0$$

$$u_i^* g_i(x^*) = 0 \quad i = 1, \dots, m$$

$$u_i^* \geq 0 \quad i = 1, \dots, m$$

A (B) feladat esetén a Kuhn-Tucker féle elsőrendű regularitási feltételek teljesülésének elégséges feltétele, hogy a $\nabla g_i(x^*)$, $i \in I(x^*)$, $\nabla_j h(x^*)$, $j=1, \dots, p$ vektorok lineárisan függetlenek legyenek. A Lagrange függvényt most az

$$/2.3/ \quad L(x, u, w) = f(x) - \sum_{i=1}^m u_i g_i(x) + \sum_{j=1}^p w_j h_j(x)$$

képlettel definiáljuk. A w itt egy p dimenziós vektor, amelynek komponensei a w_j -k. A w_j Lagrange-szorzókra nincs előjelmegkötés. Igaz a következő

3. Tétel. Ha x^* a (B) feladat megoldása és a $\nabla g_i(x^*)$, $i \in I(x^*)$, $\nabla_j h_j(x^*)$, $j=1, \dots, p$ vektorok függetlenek, akkor létezenek olyan u_i^* , w_j^* Lagrange-szorzók, hogy teljesülnek a

$$\begin{aligned} /2.4/ \quad & \nabla_x L(x^*, u^*, w^*) = 0 \\ & h_j(x^*) = 0 \quad j = 1, \dots, p \\ & u_i^* g_i(x^*) = 0 \quad i = 1, \dots, m \\ & u_i^* \geq 0 \end{aligned}$$

feltételek.

Főleg nem-konvex feladatok esetén hasznos másodrendű optimalitási kritériumok bevezetése. Nem-konvex feladatot kapunk, ha a feltételek között egyenlőségek is szerepelnek és ezek nem lineáris egyenlőségek. A másodrendű optimalitási feltétel teljesüléséhez szükség van egy másodrendű regularitási feltétel megfogalmazásához /ld. [9] /. Ehelyett egy többet követelő, de egyszerű feltétellel élünk:

/2.5/ az f, g_i függvények kétszer folytonosan differenciálhatók és a

$$\nabla g_i(x^*), \quad i \in I(x^*), \quad h_j(x^*), \quad j = 1, \dots, p$$

vektorok lineárisan függetlenek.

Az optimalitás másodrendű szükséges kritériumát fogalmazza meg a következő tétel.

4. Tétel. Ha x^* a (B) feladat megoldása és teljesüljenek a /2.5/ feltételek, akkor léteznek olyan u_i^* , w_j^* Lagrange-szorzók, amelyekkel teljesül a /2.4/ feltételrendszer, továbbá bármely $y \in E^n$ vektorra, amely kielégíti az

$$\begin{aligned} & y' \cdot \nabla g_i(x^*) = 0 && i \in I(x^*) \\ /2.6/ & y' \cdot \nabla h_j(x^*) = 0 && j = 1, \dots, p \end{aligned}$$

feltételrendszert, teljesül az

$$/2.7/ \quad y' \cdot \nabla_{xx}^2 L(x^*, u^*, w^*) y \geq 0$$

egyenlőtlenség.

A másodrendű optimalitási feltétel erősebb formája már elégséges feltétel. Ezt mondja ki az 5. Tétel.

5. Tétel. Legyen x^* a (B) feladat egy megengedett pontja, amelyre teljesülnek a /2.4/, /2.5/ feltételek. Tegyük fel továbbá, hogy ha $y \neq 0$ olyan vektor, hogy

$$\begin{aligned} y' \cdot \nabla g_i(x^*) &= 0 && \text{minden } i\text{-re, melyre } u_i^* > 0 \\ y' \cdot \nabla g_i(x^*) &\geq 0 && i \in I(x^*) \text{-ra.} \end{aligned}$$

és

$$y' \cdot \nabla h_j(x^*) = 0 \quad j = 1, \dots, p\text{-re}$$

akkor igaz az

$$/2.8/ \quad y' \cdot \nabla_{xx}^2 L(x^*, u^*, w^*) y > 0$$

egyenlőtlenség. Ekkor x^* az f függvénynek az S megengedett tartományon vett izolált lokális minimuma.

3. A SUMT módszer és az optimalitási kritériumok összefüggése.

A SUMT módszer egy érdekes levezetését adják Fiacco és McCormick az optimalitási kritériumok alapján. A SUMT módszerrel kapcsolatos vizsgálatok gyakran ebből az újszerű értelmezésből indulnak ki. Tekintsük az /A/ feladathoz tartozó feltételrendszerben a komplementaritási feltételeket. Ezek jobboldalára írjuk 0 helyett r -et, ahol $r > 0$. Az így perturbált feltételrendszer tehát a következő:

$$/3.1/ \quad \nabla f(x) - \sum_{i=1}^m u_i \nabla g_i(x) = 0 .$$

$$/3.2/ \quad u_i g_i(x) = r \quad i = 1, \dots, m$$

$$u_i \geq 0 .$$

Tegyük fel, hogy ez az egyenletrendszer x -re és u -ra megoldható, a megoldásokat jelölje $x(r)$ és $u(r)$.

Tegyük még fel, hogy $x(r)$ megengedett pont. /3.2/ -ből u_i kifejezhető s így kapjuk a

$$\nabla f(x(r)) - \sum_{i=1}^m \frac{r}{g_i(x(r))} \nabla g_i(x(r)) = 0$$

egyenletet.

A baloldal itt nem más, mint a logaritmikus büntetőfüggvény gradiense az $x(r)$ helyen. Később megfogalmazandó feltételek mellett megmutatható, hogy $x(r)$ -ben $\nabla_{xx}^2 P(x,r)$ pozitív definit elég kicsiny r esetén, tehát $x(r)$ a $P(x,r)$ függvény egy lokális minimumhelye. /ld. Fiacco-, McCormick, 3. fejezetét/

A négyzetes büntetőfüggvény módszere is levezethető a Kuhn-Tucker féle feltételek egy másik perturbációjából. Tekintsük a (B) feladatot abban a formájában, amikor minden feltétel

egyenlőség. A /2.4/ feltételrendszer helyett vezessük be a következő feltételrendszert

$$/3.3/ \quad \nabla f(x) + \sum_{j=1}^m w_j \nabla h_j(x) = 0$$

$$/3.4/ \quad w_j r = h_j \quad j=1, \dots, p .$$

Tegyük fel, hogy a /3.3/ /3.4/ feltételrendszernek van egy $x(r)$, $w(r)$ megoldása. /3.4/-ből w_j -t kifejezve és /3.3/-ba helyettesítve a

$$\nabla f(x(r)) + \sum_{j=1}^p \frac{h_j(x(r))}{r} \nabla h_j(x(r)) = 0$$

egyenletet kapjuk. A baloldal itt nem más, mint a négyzetes büntetőfüggvény gradiense az $x(r)$ helyen. Bizonyos feltételek mellett megmutatható, hogy $x(r)$ -ben $\nabla_{xx}^2 P(x,r)$ pozitív definit ha r elég kicsiny és így $x(r)$ a $P(x,r)$ függvény lokális minimuma /ld. Fiacco - McCormick, 4. fejezet/.

A Kuhn-Tucker féle elsőrendű optimalitási kritériumban szereplő Lagrange-szorzók a SUMT módszer alkalmazásakor automatikusan adódnak. Ezt a tényt pontosabban a logaritmikus büntetőfüggvény esetén fogalmazzuk meg.

6. Tétel. Legyen x^* az (A) feladat megoldása.

Az f, g_i $i = 1, \dots, m$ függvények legyenek folytonosan differenciálhatók és a $\nabla g_i(x^*)$ $i \in I(x^*)$ vektorok legyenek lineárisan függetlenek. A logaritmikus büntetőfüggvény izolált lokális minimumainak egy x^* -hoz konvergáló sorozatát jelölje $x(r)$.

Vezessük be az

$$/3.5/ \quad u_i(r) = \frac{r}{g_i(x(r))} \quad i = 1, \dots, m$$

jelölést.

Ekkor $r \rightarrow 0$ esetén $u_i(r) \rightarrow u_i^x$ ahol u_i^x az x^x ponthoz tartozó Lagrange-szorzókat jelölik.

A tétel bizonyítása egyszerű /ld. Fiacco és McCormick/.

4. Néhány mátrixelméleti eredmény.

Szükségünk lesz néhány mátrixelméleti eredményre, ezeket most foglaljuk össze. Az első eredmény Fisler-Lemma néven ismert, amelyet a 7. tételben fogalmazunk meg.

7. Tétel. Adott az n -dimenziós euklidesi tér egy lineáris altere, amelyet az

$$/4.1/ \quad a_i'x = 0 \quad i = 1, \dots, m$$

feltételrendszer definiálás egy olyan C szimmetrikus mátrix, hogy bármely /4.1/ -nek eleget tevő $x \neq 0$ vektorra

$$/4.2/ \quad x'Cx > 0$$

Legyen adott továbbá egy D szimmetrikus mátrix, amelyre teljesülnek a következő feltételek: ha x kielégíti /4.1/-et, akkor

$$/4.3/ \quad Dx = 0 .$$

Továbbá bármely

$$/4.4/ \quad y = \sum_{i=1}^m \lambda_i a_i \neq 0 \quad \text{vektorra}$$

$$/4.5/ \quad y'Dy > 0 .$$

Állítjuk, hogy ekkor bármely elegendő nagy k pozitív számra az

$$/4.6/ \quad F = C + kD$$

mátrix pozitív definit.

Bizonyítás: Válasszunk egy új ortogonális koordinátarendszert, amelynek koordinátavektorai a D mátrix sajátvektorai. A /4.1/ feltételek által meghatározott altér a D mátrix invariáns altére.

Mivel D szimmetrikus, az ortogonális kiegészítő altér, vagyis a /4.4/ alakú vektorokból alkotott altér is D invariáns altére. Ezért D alakja az új koordinátarendszerben

$$D = \begin{pmatrix} D_1 & 0 \\ 0 & 0 \end{pmatrix}$$

Itt D_1 egy $m \times m$ -es mátrix. Feltesszük, hogy az a_1, \dots, a_m vektorok lineárisan függetlenek, ami nem jelent lényegi megszorítást. A C mátrixot a D mátrixhoz hasonlóan particionáljuk:

$$C = \begin{pmatrix} C_1 & C_3 \\ C'_3 & C_2 \end{pmatrix}$$

Az n dimenziós tér vektorait az új koordinátarendszerben (y, z) alakban fogjuk felírni, ahol y m -dimenziós z $(n-m)$ -dimenziós vektorok.

A /4.1/ feltételeknek eleget tevő x vektorok azonosak a $(0, z)$ alakú vektorokkal, míg a komplementor altér vektorai

$(y, 0)$ alakúak. A /4.2/ feltétel tehát úgy fogalmazható, hogy a C_2 mátrix pozitív definit. A /4.5/ feltétel pedig azt jelenti, hogy a D_1 mátrix pozitív definit.

A D_1 ill. C_2 mátrix sajátértékeinek egy-egy pozitív alsó korlátja legyen μ_1 ill. μ_2 . Számítsuk ki az $F=C+kD$ mátrix által meghatározott kvadratikus forma értékét valamely $x=(y, z) \neq 0$ vektorra:

$$/4.7/ \quad x'Fx = ky' D_1 y + y' C_1 y + 2y' C_3 z + z' C_2 z.$$

A jobboldal alsó becsléséhez felhasználjuk az

$$/4.8/ \quad y' \cdot D_1 y \geq \mu_1 \|y\|^2, \quad z' C_2 z \geq \mu_2 \|z\|^2$$

egyenlőtlenségeket. Legyen γ_1, γ_3 a C_1, C_3 mátrixok normájának egy-egy felső becslése. Ekkor kapjuk az

$$/4.9/ \quad y' C_1 y \leq -\gamma_1 \|y\|^2, \quad y' C_3 z \leq -\gamma_3 \|y\| \|z\|$$

egyenlőtlenségeket. Vezessük be az $y = \|y\|$, $z = \|z\|$ jelöléseket. A /4.8/, /4.9/ becsléseket felhasználva a következő alsó becslést kapjuk:

$$x'Fx \geq (k\mu_1 - \gamma_1) y^2 - 2\gamma_3 yz + \mu_2 z^2.$$

A jobboldali kifejezés egy kétdimenziós kvadratikus forma, amelynek mátrixa

$$\begin{pmatrix} k\mu_1 - \gamma_1 & -\gamma_3 \\ -\gamma_3 & \mu_2 \end{pmatrix}$$

Világos, hogy ez a kvadratikus forma pozitív definit, ha csak k elég nagy, s ezért pozitív definit az $x'Fx$ kvadratikus forma is. Ezzel a 7. tételt bebizonyítottuk.

Egyszerű, de hasznos tétel a következő

8. Tétel. Legyen adott egy C $n \times n$ -es szimmetrikus mátrix és egy A $n \times m$ -es mátrix. Az A mátrix rangja legyen m és teljesüljön a következő feltétel: ha valamely n dimenziós $x \neq 0$ vektorra

$$/4.11/ \quad x'A = 0$$

akkor

$$/4.12/ \quad x'Cx > 0$$

Állítás, hogy a

$$/4.13/ \quad G = \begin{pmatrix} C & A \\ A' & 0 \end{pmatrix}$$

mátrix nem szinguláris.

Bizonyítás: Tekintsünk egy $n+m$ -dimenziós

$$W = \begin{pmatrix} x \\ v \end{pmatrix}$$

vektort, melyre

$$/4.14/ \quad Gw = 0.$$

Részletesebben írva kapjuk a

$$/4.15/ \quad Cx + Av = 0$$

$$A'x = 0$$

egyenlőségrendszer. Szorozzuk be az első egyenlőséget az x' vektorral balról. A második egyenlet figyelembevételével az

$$x' C x = 0$$

egyenletet kapjuk. Mivel x eleget tesz a /4.11/ feltételnek, következik, hogy $x = 0$. Az /4.15/ egyenletrendszer első egyenlete tehát

$$A v = 0$$

-ra egyszerűsödik. Mivel A rangja m , következik, hogy $v = 0$. Tehát a /4.1.4/ egyenlet bármely w megoldására $w = 0$.

A SUMT módszerrel kapcsolatban kifejlesztett extrapolációs technika alkalmazásával hasznos segédeszköznek fog bizonyulni a Householder-triangularizáció, amelyet most ismertetünk. Householder-transzformációnak nevezünk egy

$$/4.16/ \quad P(u) = I - \beta u u'$$

alaku mátrixszal megadott transzformációt, ahol I az egységmátrix és $\beta = 2/u'u$. A Householder transzformáció úgy értelmezhető, mint egy az u normálvektorral meghatározott síkra való tükrözés. Ezt a következőképpen láthatjuk be. Tetszőleges x vektort állítsunk elő az

$$x = \lambda u + v$$

alakban, ahol v ortogonális u -ra. Alkalmazzuk x -re a $P(u)$ transzformációt:

$$/4.17/ \quad P(u)x = \lambda u + v - \lambda \beta u u' u = \lambda u + v - 2\lambda u = -\lambda u + v.$$

Látható, hogy $P(u)$ valóban tükrözés. Ebből következik az is, hogy a Householder-transzformáció ortogonális transzformáció.

Megmutatjuk, hogy bármely x vektorhoz található olyan u , hogy a $P(u)x$ vektornak egyetlen nem-zéró komponense van előre kitüntetett helyen. Valóban legyen a kitüntetett koordináta az i -edik. Nyujtsuk meg az i -dik koordinátavektort az x vektorral egyenlő hosszúságúra, a kapott vektort jelölje f_i . Mivel $P(u)$ független a keresett u hosszától, csupán annak irányától függ, feltehetjük, hogy x az

$$/4.18/ \quad x = u - v$$

alakban írható, ahol v ortogonális u -ra. Ekkor a megkívánt $P(u)x = f_i$ egyenlőség /4.17/ alapján felírható az

$$/4.19/ \quad f_i = -u + v$$

alakban. A /4.18/ és /4.19/ egyenletekből u, v egyértelműen kiszámítható.

A Householder transzformációk segítségével valósítjuk meg az un . Householder-triangularizációt, amit szokás QR felbontásának is nevezni. Legyen A egy $m \times n$ -es mátrix, ahol $m \leq n$. Szorozzuk meg A -t balról egy $P(u_1)$ mátrixszal úgy, hogy az

$$A_1 = P(u_1)A$$

mátrix első oszlopában az első elem kivételével valamennyi elem 0 legyen. A_1 tehát a következő alakú

$$A_1 = \begin{pmatrix} \bullet & \square \\ 0 & \square \end{pmatrix}$$

Sorozzuk meg balról A_1 -et egy olyan $P(u_2)$ mátrixszal, amely az első koordinátavektor által kifeszített altéren identitás, a maradék koordinátavektorok által kifeszített $n-1$ dimenziós altéren pedig úgy hat, hogy az $A_2 = P(u_2)A_1$ mátrix második oszlopában a második elem után csupa 0 áll.

Ezt az eljárást folytatjuk, amíg lehet. Előfordulhat, hogy a soronkövetkező oszlopban a diagonálem alatt már valamennyi elem 0, de a később következő oszlopokban ez még nem áll fenn. Ilyenkor az oszlopokat permutáljuk. A k-adik lépés általános alakja tehát

$$/4.20/ \quad A_{k+1} = P(u_{k+1}) A_k \tilde{\pi}_{k+1} ,$$

ahol $P(u_{k+1})$ egy Householder-transzformáció, $\tilde{\pi}_{k+1}$ egy permutációmátrix.

Ha A rangja m , akkor az eljárás m lépés után ér véget és eredményül egy

$$/4.21/ \quad A_m = QA\tilde{\pi}$$

mátrixot kapunk, ahol

$$Q = P(u_m) \dots P(u_1)$$

és

$$\tilde{\pi} = \tilde{\pi}_1 \dots \tilde{\pi}_m .$$

A Q mátrix ortogonális. Az A_m mátrix alakja

$$A_m = \begin{pmatrix} \begin{array}{c|c} \square & \\ \hline 0 & \end{array} \\ \hline 0 \end{pmatrix}$$

(The square contains a diagonal line and the letter R)

ahol R felső-háromszög alakú nem-szinguláris mátrix.
/ld. [2] , [12] /

5. Konvexitási kérdések

A nemlineáris programozási feladatok között különleges helyet foglalnak el a konvex feladatok. Az (A) alakban megfogalmazott feladatot konvexnek nevezzük, ha az $f, -g_i$ függvények konvex függvények. A konvex feladatok jelentőségét az adja, hogy ezen feladatok esetén a Kuhn-Tucker féle feltétel az optimalitásnak elégséges feltétele is. Ezt mondja ki a következő

9. Tétel. Tekintsük az (A) feladatot. Legyenek az f, g_i függvények folytonosan differenciálható konvex függvények, melyek értelmezve vannak egy az R megengedett tartományt tartalmazó konvex nyílt halmazon. Az $x^* \in R$ pontban teljesülnek a (2.1) optimalitási feltételek. Ekkor x^* az (A) feladat megoldása.

A tétel és annak bizonyítása megtalálható [27] -ban. A tétel kiterjeszthető kvázikonvex függvényekre is. Az idevontakozó első eredmények Mangasariantól valók. Ujabbban Ferland bizonyította a 9. tétel egy általánosítását. Egy újabb általánosítást adunk a 11. tételben, amely különbözik az említett szerzők eredményeitől. Előbb azonban a kvázikonvex függvények egy jellemzését adjuk.

Megmutatjuk, hogy kvázikonvex függvények egy igen tág osztálya alkalmas monoton transzformáció segítségével konvex függvénybe megy át. Előbb azonban bevezetjük a szigorúan kvázikonvex függvény fogalmát. Egy D konvex nyílt halmazon kétszer folytonosan differenciálható f kvázikonvex függvényt szigorúan kvázikonvexnek mondunk ha f nivófelületeinek görbülete minden $x \in D$ -re zérustól különböző. Igaz a következő

10. Tétel. Legyen $f(x)$ egy kétszer folytonosan differenciálható szigorúan kvázikonvex függvény, mely értelmezve van egy D konvex nyílt halmazon. Legyen D_0 egy D -ben fekvő konvex

kompakt halmaz. Ekkor létezik olyan k_0 szám, hogy $k > k_0$ esetén az

$$/5.1/ \quad F(x) = e^{kf(x)}$$

függvény konvex D_{0-n} .

Bizonyítás: Legyen x_0 a D halmaz egy pontja és tekintsük a

$$/5.1/ \quad P = \{x_0 + v \mid \nabla'_x f(x_0) \cdot v = 0\}$$

egyenlettel meghatározott hipersíkot. Ennek a hipersíknak a paraméteres előállítását legyen a következő

$$P = \{x \mid x = By + x_0\}$$

ahol B egy $n \times (n-1)$ -es mátrix, y pedig tetszőleges $(n-1)$ dimenziós vektor. Mivel $f(x)$ kvázikonvex, azért az $f(x)$ függvénynek a P -re való megszorítása x_0 -ban minimális. Ez úgy is megfogalmazható, hogy a

$$/5.2/ \quad g(y) = f(x_0 + By)$$

függvény $y = 0$ -ban éri el minimumát. Ezért a $\nabla_{yy}^2 g(y)$ mátrix pozitív szemidefinit. Az y vektorok paraméterezik az $f(x)$ nivófelületéhez vont P érintősíkot. A nivófelületnek különböző irányokban vett görbületeit az

$$y' \nabla_{yy}^2 (g(0)) y$$

kifejezések adják. Mivel $f(x)$ szigorúan kvázikonvex ez a kifejezés $y \neq 0$ esetén nem lehet 0, tehát $\nabla_{yy}^2 g$ pozitív definit. Ez /5.2/ felhasználásával úgy is írható, hogy minden $y \neq 0$ vektorra fennáll az

$$/5.3/ \quad y' B' \nabla_{xx}^2 f(x) B y > 0$$

egyenlőtlenség. Vagyis a $v' \nabla_{xx}^2 f(x) v$ kvadratikus alak pozitív definit a

$$/5.4/ \quad \nabla' f(x_0) \cdot v = 0$$

egyenlet által meghatározott altéren.

Számítsuk most ki az $F(x) = e^{kf(x)}$ függvény Hesse-mátrixát:

$$\nabla_x F(x) = k e^{kf(x)} \nabla f(x)$$

/5.5/

$$\nabla_{xx}^2 F(x) = k e^{kf(x)} \left(\nabla_{xx}^2 f(x) + k \nabla f(x) \nabla' f(x) \right)$$

A

$$C = \nabla_{xx}^2 f(x)$$

$$D = \nabla f(x) \nabla' f(x)$$

$$a_1 = \nabla f(x)$$

választással teljesülnek a Finsler-lemma feltételei, ezért az /5.5/ jobboldalán álló mátrix pozitív definit, hacsak k elég nagy. Ezzel a 10. tételt bebizonyítottuk.

A most bebizonyított tétel szoros kapcsolatban van a kvázikonvex kvadratikus függvények Kéri Gerzson által adott jellemzésével.

A 10. tétel alapján megmutatható, hogy a Kuhn-Tucker feltétel teljesülése szigorúan kvázikonvex feladat esetén is az optimalitás elégséges feltétele. Pontosabban igaz a következő:

11. Tétel. Legyenek az (A) feladatban szereplő $f, -g_i$ függvények kétszer folytonosan differenciálható szigorúan kvázikonvex függvények, amelyek értelmezve vannak egy az R megengedett tartományt tartalmazó D konvex halmazon. Az R, D halmazok legyenek korlátosak. Az $x^* \in R$ pontban teljesülnek a (2.1) Kuhn-Tucker feltételek. Ekkor x^* az (A) feladat megoldása.

Bizonyítás: A 10. tétel szerint létezik oly k szám, hogy az

$$F(x) = e^{kf(x)}$$
$$G_i(x) = -e^{-kg_i(x)} + 1$$

függvények konvex ill. konkáv függvények, melyek értelmezve vannak egy R -et tartalmazó konvex nyílt D_0 halmazon. Az eredeti feladattal ekvivalens a következő feladat

$$\begin{aligned} /5.6/ \quad & \min F(x) \\ & G_i(x) \geq 0 \quad i = 1, \dots, m. \end{aligned}$$

Mivel $\nabla F(x) = \text{skalár} \cdot \nabla f(x)$

és

$$\nabla G_i(x) = \text{skalár} \cdot \nabla g_i(x)$$

következik, hogy x^* az /5.6/ feladatnak is Kuhn-Tucker pontja. A 9. tétel alapján tehát x^* az /5.6/ feladatnak és így a vele ekvivalens eredeti (A) feladatnak is megoldása.

Konvex feladatok esetén a logaritmikus, a reciprok, valamint a négyzetes büntetőfüggvény alkalmazása esetén a $P(x,r)$ segédfüggvény konvex. Ezért $P(x,r)$ globális minimumának a meghatározása számos ismert eljárással lehetséges. Sajnos kvázikonvex feladat esetén hasonló állítás nem érvényes. Ha azonban a célfüggvény konvex, a feltételi függvények pedig kvázi-

konkávak, akkor a 10. Tétel alapján várható, hogy a

$$P(x,r) = f(x) + \frac{1}{r} \sum e^{-\frac{1}{r} g_i(x)}$$

exponenciális büntetőfüggvény konvex lesz. Speciális esetben a logaritmikus büntetőfüggvény is konvex segédfüggvényhez vezet. Ez a helyzet a sztohasztikus programozás bizonyos feladataiban, ahol a feltételekben szereplő függvények logaritmikusan konkáv függvények.

A kvázikonvex programozási feladatokkal kapcsolatban sok szép eredmény született Magyarországon. Martos Béla és Kéri Gerzson munkái főleg elméleti eredményeket tartalmaznak, /ld. [24], [20]/. Kovács László Béla a gradiensvetítési módszertm Prékopa András pedig a Zoutendijk-féle megengedett irányok módszerét terjesztette ki kvázikonvex programozási feladatokra /ld. [21], [30]/. Ez utóbbi két eredmény lényegében a 10. Tételből egyszerűen levezethető.

II. AZ EXTRAPOLÁCIÓS TECHNIKA ÉS KITERJESZTÉSE

1. Az extrapolációs technika megalapozása

Az extrapoláció alkalmazása azt a célt szolgálja, hogy a SUMT módszer konvergenciáját meggyorsítsuk. Az extrapolációs technika elméleti hátterét először a logaritmikus büntetőfüggvény esetén mutatjuk be. Legyen x^* az (A) feladat egy Kuhn-Tucker pontja.

A logaritmikus büntetőfüggvényt jelölje $P(x,r)$ tehát

$$/1.1/ \quad P(x,r) = f(x) - r \sum_{i=1}^m \ln g_i(x)$$

Arra akarunk választ kapni, milyen feltételek mellett létezik egy x^* -hoz konvergáló folytonos $x(r)$ görbe, ahol $x(r)$ minden $r > 0$ mellett a $P(x,r)$ függvény izolált lokális minimuma.

Előkészítésképpen bebizonyítjuk a következő tételt.

12. Tétel Legyenek az f, g_i függvények kétszer folytonosan differenciálhatók. Az x^* pontban teljesüljenek a másodrendű elégséges feltételek, teljesüljön a szigorú komplementaritási feltétel és a $\nabla g_i(x^*)$ $i \in J(x^*)$ vektorok legyenek lineárisan függetlenek.

Ekkor, ha $x(r)$ a $P(x,r)$ függvénynek egy x^* -hoz elég közel eső stacionárius pontja, és r elég kicsiny, akkor $\nabla_{xx}^2 P(x(r), r)$ pozitív definit.

Bizonyítás: Vezessük be az

$$/1.2/ \quad u_i(r) = \frac{r}{g_i(x(r))} \quad i = 1, \dots, m,$$

jelölést. Az x^x ponthoz tartozó Lagrange-szorzókat jelölje: u_i^x . A $\nabla P(x(r), r) = 0$ egyenlőségből és a $\nabla g_i(x^x)$, $i \in I(x^x)$ vektorok lineáris függetlenségéből következik, hogy $u_i(r)$ és u_i^x eltérése tetszőlegesen kicsiny, ha csak $x(r)$ és x^x eltérése elég kicsiny.

Írjuk fel a $P(x, r)$ függvény Hesse-mátrixát:

$$\nabla^2 P(x, r) = \nabla^2 f(x) + r \nabla^2 I(x).$$

Itt

$$\nabla^2 I(x) = - \sum_{i=1}^m \frac{\nabla^2 g_i(x)}{g_i(x)} + \sum_{i=1}^m \frac{\nabla g_i(x) \cdot \nabla g_i'(x)}{g_i^2(x)}.$$

Az $x = x(r)$ helyen helyettesítsük be az $u_i(r)$ -ek fenti értékeit. Ekkor a

$$/1.3/ \quad \nabla^2 P(x(r), r) = \nabla^2 f(x(r)) - \sum_{i=1}^m u_i(r) \nabla^2 g_i(x(r)) + \\ + r^{-1} \sum_{i=1}^m u_i^2(r) \nabla g_i(x(r)) \cdot \nabla g_i'(x(r))$$

kifejezést kapjuk. Vezessük be a

$$/1.4/ \quad \Gamma(x(r)) = \sum_{i=1}^m u_i^2(r) \nabla g_i(x(r)) \cdot \nabla g_i'(x(r))$$

jelölést. A Lagrange-függvényt jelölje $L(x, u)$.

Az /1.4/ képlet jobboldalából a

$$/1.5/ \nabla^2 P(x(r), r) = \nabla^2 L(x(r)) + r^{-1} \Gamma(x(r))$$

kifejezést kapjuk.

Rögzített r esetén a jobboldalon álló mátrixnak a

$$/1.6/ \nabla^2 L(x^x, u^x) + r^{-1} \Gamma(x^x)$$

mátrixtól való eltérése tetszőlegesen kicsiny, hacsak $x(r)$ elég közel van x^x -hez.

A 7. tétel alapján tudjuk, hogy az /1.6/ mátrix pozitív definit, hacsak $r < r_0$, ahol r_0 egy elég kicsiny szám. Az említett tétel bizonyításából az is világos, hogy $r < r_0$ esetén az /1.6/ mátrix sajátértékeire egy r -től függetlenül pozitív alsó korlát adható. Ezért az /1.6/ mátrixtól kicsinyt különböző /1.5/ mátrix is pozitív definit, amint állítottuk.

A 12. tétel felhasználásával könnyen bizonyítható a következő tétel.

13. Tétel Legyenek az f, g_i függvények kétszer folytonosan differenciálhatók. Az x^x pontban teljesüljenek a másodrendű elégséges feltételek, teljesüljön a szigorú komplementaritási feltétel, és a $\nabla g_i(x^x), i \in I(x^x)$ vektorok legyenek lineárisan függetlenek. Ekkor létezik a $P(x, r)$ függvény feltétel nélküli izolált lokális minimumainak egy folytonos trajektóriája, amely tart x^x -hoz.

A 13. tétel lényegében megtalálható a Fiacco és McCormick könyvében. Az ott közölt bizonyítás hibás, azonban könnyen kijavítható.

A következőkben $x(r)$ -rel mindig egy a 13. tételben szereplő trajektóriát jelölünk, $r=0$ esetén legyen $x(r) = x^*$. Az extrapoláció alkalmazásának lehetőségét biztosítja a következő

14. Tétel. A 13. tétel feltételei mellett az $x(r)$ görbe folytonosan differenciálható $r \geq 0$ esetén.

A 14. Tétel egy közvetlen bizonyítása az $n+m$ egyenletből álló

$$/1.7/ \quad \nabla f(x(r)) - \sum_{i=1}^m u_i(r) \nabla g_i(x(r)) = 0$$

$$u_i(r) g_i(x(r)) = r$$

nemlineáris egyenletrendszer vizsgálatán alapul. Az egyenletrendszer r szerinti deriválásával a dx/dr és du/dr értékekre egy lineáris egyenletrendszert kapunk:

$$/1.8/ \quad \begin{pmatrix} \nabla^2 L(x(r), u(r)), & \nabla g_1, & \dots, & \nabla g_m \\ u_1 \nabla' g_1 & g_1 & & \\ \vdots & & \ddots & \\ u_m \nabla' g_m & & & g_m \end{pmatrix} \begin{pmatrix} \frac{dx}{dr} \\ \frac{du}{dr} \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ \vdots \\ 1 \end{pmatrix} \left. \begin{array}{l} \} \\ \} \\ \} \\ \} \end{array} \right\} \begin{array}{l} n \\ \\ \\ m \end{array}$$

Ennek az egyenletrendszernek a vizsgálata elméletileg vonzó, azonban számítási célokra nem a legalkalmasabb. Ezt hangsúlyozza Fiacco is [9]. Ezért később egy új vizsgálati módszert vezetünk be.

A 14. tétel alapján a SUMT módszer konvergenciáját gyorsíthatjuk. Az x^* pontból kiindulva az $x(r)$ görbét lineáris függvény görbéjével közelíthetjük:

$$/1.9/ \quad x(r) \approx x^x + rh_0$$

ahol

$$/1.10/ \quad h_0 = \left. \frac{dx(r)}{dr} \right|_{r=0}$$

A h_0 vektort közelítőleg úgy határozzuk meg, hogy kiszámítjuk $x(r)$ -et két különböző r értékre, és utána az /1.9/ formula alapján r -et kiküszöböljük.

Vezessük be az

$$/1.11/ \quad x^1 = x(r_1) \quad , \quad x^2 = x(r_2)$$

jelöléseket. A mondott eljárással kapjuk a következő becslést

$$/1.12/ \quad h_0 \approx (x^1 - x^2) / (r_1 - r_2) .$$

Az x^x becslése pedig

$$/1.13/ \quad x^x \approx x^1 - h_0 r_1 \approx (r_1 x^2 - r_2 x^1) / (r_1 - r_2) .$$

Megjegyezzük, hogy az extrapolációs technikát alkalmazhatjuk $x(r)$ becslésre is, ha r értéke kicsiny. Ha az f, g_i függvények magasabbrendben differenciálhatók, akkor az $x(r)$ görbét másodfoku görbével is közelíthetjük. Így olyan extrapolációs formulát kapunk, amelynek a hibája r^3 nagyságrendű.

2. Az extrapolációs technika kiterjesztése.

Az extrapolációs technikát most olyan pontokra is ki fogjuk terjeszteni, amelyek nem azonosak valamely $x(r)$ ponttal.

Induljunk ki a

$$/2.1/ \quad \nabla P(x,r) = \nabla f(x) + r \nabla I(x) = 0$$

egyenletből. Ez az egyenlet az $x(r)$ görbét definiálja. Ezt a görbét beágyazzuk egy $x(r,\theta)$ görbeseregbe, ahol θ egy n -dimenziós paraméter. A görbesereg egyenletét a /2.1/ egyenletből pertubációval származtatjuk:

$$/2.2/ \quad \nabla f(x) + r \nabla I(x) + r\theta = 0.$$

A /2.2/ egyenlet baloldala nem más, mint a

$$/2.3/ \quad P(x,r,\theta) = f(x) + rI(x) + r\theta'x$$

függvény gradiense. Ha az (A) feladatban $f(x)$ helyett az $f(x) + r\theta'x$ függvényt vesszük célfüggvénynek, akkor $P(x,r,\theta)$ a módosított feladat logaritmikus büntetőfüggvénye. A módosított feladatra is fennállnak a 12. tételben megfogalmazott feltételek, ha ezek az eredeti feladatra fennállnak, s ezért ha $x(r,\theta)$ a /2.3/ egyenletnek egy x^* -hoz elég közelső megoldása, akkor $x(r,\theta)$ a $P(x,r,\theta)$ függvény lokális minimuma. A 13. tétel analógiájára igaz a

15. Tétel. A 13. tétel feltételei mellett létezik a $P(x,r,\theta)$ függvény izolált lokális minimumainak egy folytonos trajektóriája, amely tart x^* -hoz, amint $r \rightarrow 0$, bármely rögzített θ mellett.

A 14. tételhez hasonlóan bizonyítható a

16. Tétel. Tekintsük a /2.3/ egyenlet által meghatározott $x(r,\theta)$ trajektóriákat, amikor θ egy korlátos halmazból

való. Az $x^x(r, \theta)$ trajektória minden $r_0 \cong r \cong 0$ érték mellett folytonosan differenciálható r szerint.

A fenti beágyazási módszert algoritmusban a következőképpen alkalmazhatjuk. Az $x(r, \theta)$ görbék kitöltik a megengedett tartomány egy részét. Adott $x \in R^0$ ponthoz keressünk olyan $x(r, \theta)$ görbét, amely áthalad x -en. Ehhez meg kell oldanunk a /2.2/ egyenletet, amelyben az ismeretlenek most r, θ .

Mivel a /2.2/ egyenletrendszer n db. skaláregyenletből áll, az ismeretlenek száma pedig $n+1$, bevezetünk még egy egyenletet. Egy egyszerű egyenlet lehet a következő

$$/2.4/ \quad \nabla' f(x) \cdot \theta = 0 .$$

Az r, θ meghatározása most már könnyű feladat. Szorozzuk be /2.2/-t $\nabla f(x)$ -szel. Így egy egyenletet kapunk, r -re, amelynek a megoldása

$$/2.5/ \quad r = r(x) = - \frac{\nabla' f(x) \nabla f(x)}{\nabla' I(x) \nabla f(x)} .$$

A θ paraméter értéke is könnyen meghatározható, de erre a továbbiakban nem lesz szükség.

Ezekután meghatározzuk az $x(r, \theta)$ görbe érintővektorát. A /2.2/ egyenletet differenciáljuk r szerint:

$$\left(\nabla^2 f(x) + r \nabla^2 I(x) \right) \frac{dx}{dr} + \nabla I + \theta = 0 .$$

A /2.1/ egyenlet felhasználásával ez így is írható

$$/2.6/ \quad \left(\nabla^2 f(x) + r \nabla^2 I(x) \right) \frac{dx}{dr} = r^{-1} \nabla f .$$

A 12. tétel alapján a baloldalon álló együttható mátrix pozitív definit, hacsak r elég kicsiny és θ egy adott korlátos halmazból való. Így a /2.6/ egyenlet egyértelműen megoldható. A megoldást jelölje h .

Mivel az $x(r, \theta)$ görbe folytonosan differenciálható $r \geq 0$ esetén az x^* megoldás egy jó közelítését adja az

$$/2.7/ \quad x^* \approx x - rh$$

képlet.

Az extrapolációs technika fenti kiterjesztésével kapcsolatban láttuk, hogy egy adott x ponton át több $x(r, \theta)$ görbe fektethető. Kérdés, hogy r ill. θ választása milyen hatással van a közelítés pontosságára. Világos, hogy r egyértelműen meghatározza θ -t. Az $x(r, \theta)$ görbe érintővektora az x pontban ezért csak r függvénye. Jelöljük ezt $h(r)$ -rel. Legyen az f függvény lineáris. Ekkor a /2.6/ egyenlet így egyszerűsödik:

$$/2.7/ \quad r^2 \nabla^2 I(x) \frac{dx}{dr} = \nabla f .$$

Vagyis a $h(r)$ iránya független r -től. Célszerű tehát a feladatokat olyan alakra hozni, hogy $f(x)$ lineáris függvény legyen. Egy ismert transzformáció a következő:

$$/2.8/ \quad \begin{aligned} & \min u \\ & u - f(x) \geq 0 \\ & g_i(x) \geq 0 \quad i = 1, \dots, m . \end{aligned}$$

3. A Householder-triangularizáció alkalmazása.

Az extrapolációs technika kiterjesztésénél nyitott kérdés, meg tudjuk-e oldani hatékonyan a /2.6/ egyenletet. A következőkben ezzel a kérdéssel foglalkozunk. Az egyszerűség kedvéért a $\theta = 0$ esettel foglalkozunk.

Kiindulópontunk a

$$/3.1/ \quad \nabla P(\mathbf{x}, r) = \nabla f(\mathbf{x}) + r \nabla I(\mathbf{x}) = 0 \quad 0$$

egyenlőség, amely érvényes az $\mathbf{x}(r)$ görbe mentén.

Számítsuk ki $\frac{d\mathbf{x}(r, \theta)}{dr}$ -et. A /3.1/ egyenlet r szerinti deriválásából kapjuk a

$$/3.2/ \quad (\nabla^2 f(\mathbf{x}) + r \nabla^2 I(\mathbf{x})) \frac{d\mathbf{x}}{dr} + \nabla f(\mathbf{x}) = 0$$

egyenletet. Vezessük be az

$$/3.3/ \quad u_i(r) = \frac{r}{g_i(\mathbf{x}(r, \theta))}$$

és

$$/3.4/ \quad \Gamma = \Gamma(\mathbf{x}(r)) = \sum_{i=1}^m u_i^2 \nabla g_i(\mathbf{x}) \nabla g_i(\mathbf{x})$$

jelöléseket. A /3.2/ egyenlet ekkor így is írható:

$$/3.3/ \quad (\nabla^2 L(\mathbf{x}, u) + r^{-1} \Gamma(\mathbf{x})) \frac{d\mathbf{x}}{dr} = r^{-1} \sum_{i=1}^m u_i \nabla g_i(\mathbf{x}).$$

Szorozzuk végig r -rel és vezessük be a

$$/3.4/ \quad D = D(\mathbf{x}(r)) = r \nabla^2 L(\mathbf{x}, u) + \Gamma(\mathbf{x})$$

jelölést. Ekkor $h = \frac{dx}{dr}$ eleget tesz a

$$/3.5/ \quad D h = \sum_{i=1}^m u_i \nabla g_i(x)$$

egyenletnek.

A /3.5/ egyenlet elemzésének további alapjául az egyenlet egy alkalmas felbontása szolgál. Az egyszerűség kedvéért tegyük fel, hogy $I(x^x) = \{1, \dots, k\}$.

Mivel a $\nabla g_i(x^x)$, $i = 1, \dots, k$ vektorok lineárisan függetlenek, azért ugyanez fennáll a $\nabla g_i(x(r))$ vektorokra is, hacsak r elég kicsiny. A $h=h(r)$ vektort $h(r) = h'(r) + h''(r)$ alakban fogjuk felbontani, ahol $h'(r)$ és $h''(r)$ ortogonálisak és $h'(r)$ a $\nabla g_i(x(r))$ vektorok által kifeszített altérhez tartozik. Ezt a felbontást a Householder-triangularizáció segítségével valósítjuk meg.

Tekintsük az $u_i(r) \nabla g_i(x(r))$, $i = 1, \dots, k$ vektorokból alkotott

$$/3.6/ \quad N(r) = (u_1(r) \nabla g_1(x(r)), \dots, u_k(r) \nabla g_k(x(r)))$$

mátrixot.

Vezessük be a

$$/3.7/ \quad D^0(r) = rA(r) + N(r)N'(r)$$

mátrixot, ahol $A(r) = \nabla^2 L(x(r), u(r))$.

A $D^0(r)$ mátrix a $D(r)$ mátrixból az $u_i^2(r) \nabla g_i(x(r)) \nabla' g_i(x(r))$ alakú tagok elhagyásával adódik, ahol $i \notin I(x^x)$. A két mátrix eltérése tehát r^2 nagyságrendű. Látni fogjuk az alábbiakban, hogy ez a hiba h határértékének meghatározásánál nem játszik szerepet.

A /3.5/ egyenletben módosítjuk a jobboldalt is, amennyiben az összegezést csak $i \in I(x^x)$ -ra terjesszük ki és elhagyjuk az $r \in \theta$ tagot. Az ebből származó hiba r nagyságrendű. A módosított jobboldal felírható $N \underline{1}$ alakban is, ahol $\underline{1}$ egy csupa egyesből álló k dimenziós vektor. A /3.5/ egyenlet helyett így kapjuk a

$$/3.8/ \quad D^0(r)h^0(r) = (rA(r) + N(r)N'(r))h^0(r) = N(r)\underline{1}$$

egyenletet.

Az N mátrixra alkalmazzuk a Householder-triangularizációt. Az első fejezet 4. pontjában leírt módon meghatározunk egy $Q(r)$ ortogonális mátrixot és egy π permutációmátrixot úgy, hogy $Q(r)N(r)\pi$

$$/3.9/ \quad \begin{pmatrix} R(r) \\ 0 \end{pmatrix}$$

alaku legyen, ahol $R(r)$ egy felső-háromszög alaku nem-szinguláris mátrix.

Szorozzuk meg a /3.8/ egyenletet balról Q -val és vezessük be az

$$/3.10/ \quad l(r) = Q(r)h^0(r)$$

jelölést. Mivel $Q(r)$ ortogonális, h^0 és l között a

$$/3.11/ \quad h^0(r) = Q'(r)l(r)$$

összefüggés áll fenn.

A $Q(r)A(r)Q'(r)$ mátrixot jelöljük egyszerűen $B(r)$ -rel.
A $B(r)$ mátrixot particionáljuk a

$$/3.12/ \quad B(r) = \left(\begin{array}{cc} \overbrace{B_1(r)}^k & \overbrace{B_2(r)}^{n-k} \\ B_2'(r) & B_3(r) \end{array} \right) \left. \begin{array}{l} \} k \\ \} n-k \end{array} \right.$$

alakban. A $Q(r)N(r)N'(r)Q'(r)$ mátrix /3.9/ figyelembevételével a következő alakú

$$/3.13/ \quad k \left\{ \begin{array}{cc} \overbrace{R(r)R'(r)}^k & 0 \\ 0 & 0 \end{array} \right.$$

Végül a $Q(r)N(r)\underline{1}$ vektor a

$$/3.14/ \quad k \left\{ \begin{array}{c} R(r)\underline{1} \\ 0 \end{array} \right.$$

alakban bontható fel. Az $l(r)$ vektort is particionáljuk az

$$/3.15/ \quad l(r) = \begin{pmatrix} l_1(r) \\ l_2(r) \end{pmatrix}$$

alakban.

A jobboldalon álló vektor particionált alakja a Q -val való szorzás után

$$\begin{pmatrix} R(r)\underline{1} \\ 0 \end{pmatrix}$$

A /3.8/ egyenletet végül a következőképpen particionáljuk:

$$\begin{aligned} /3.16/ \quad rB_1(r)l_1(r) + rB_2(r)l_2(r) + R(r)R'(r)l_1(r) &= R(r)\underline{1} \\ rB_2'(r)l_1(r) + rB_3(r)l_2(r) &= 0 \end{aligned}$$

A /3.16/ egyenletrendszer megoldása igen könnyen kapható. Az első egyenletből r nagyságrendű tagok elhanyagolásával az

$$/3.17/ \quad R(r)R'(r)l_1^0(r) = R(r)\underline{1}$$

közelítő egyenletet kapjuk. Mivel $R(r)$ nem szinguláris, innen

$$/3.18/ \quad l_1^0(r) = (R'(r))^{-1} \underline{1}$$

A /3.16/ egyenletrendszer második egyenletét r -rel elosztva a

$$/3.19/ \quad B_2'(r)l_1(r) + B_3(r)l_2(r) = 0$$

egyenletet kapjuk. Megmutatjuk, hogy $B_3(r)$ nem-szinguláris $r_0 \cong r \cong 0$ esetén. Feltevésünk szerint ugyanis teljesülnek a másodrendű elégséges feltételek. Ezért minden x^x -hoz elég közel eső $x(r)$ pontban minden vektorra, amely eleget tesz a

$$/3.20/ \quad \nabla'g_1(x(r)) \cdot v = 0 \quad i = 1, \dots, k$$

feltételeknek, teljesül a

$$/3.21/ \quad v'A(r)v \cong \mu \|v\|^2$$

egyenlőtlenség, ahol $\mu > 0$. A Q transzformációval bevezetett új koordinátarendszerben a /3.20/ egyenlőség által meghatározott v vektorok a

$$v = \begin{pmatrix} 0 \\ w \end{pmatrix} \begin{matrix} \} & k \\ \} & n-k \end{matrix}$$

alakban írható. A /3.21/ feltételből pedig a

$$/3.22/ \quad w' B_3(r) w \cong \mu \| w \|^2$$

egyenlőtlenséget kapjuk. Tehát $B_3(r)$ valóban nem szinguláris $r_0 \cong r \cong 0$ esetén.

A /3.19/ egyenletben $l_1(r)$ helyett írjuk be a fentebb kiszámított $l_1^0(r)$ közelítő megoldást. Így $l_2(r)$ -re a következő közelítő értéket kapjuk:

$$/3.23/ \quad l_2^0(r) = - B_3^{-1}(r) B_2(r) l_1^0(r)$$

Az ismerttetett eljárással kapott

$$/3.24/ \quad l^0(r) = \begin{pmatrix} l_1^0(r) \\ l_2^0(r) \end{pmatrix}$$

közeliítő megoldás hibája r nagyságrendű. Ezért a $h^0(r)$ vektort közelítő

$$/3.25/ \quad h^{00}(r) = Q'(r) l^0(r)$$

vektor hibája is r nagyságrendű.

A közölt gondolatmenet alapján könnyű megmutatni azt is, hogy a /3.5/ egyenletnek a /3.8/ egyenlettel való helyettesítése a megoldások között r nagyságrendű hibát eredményez.

Ha tehát alkalmazzuk az

$$/3.26/ \quad x^* \approx x(r) - r h^{00}(r)$$

közeliítést, ennek hibája r^2 nagyságrendű.

Megjegyezzük, hogy a fenti elemzés alapján egyszerű bizonyítást kaphatunk a 14. Tételre is. Ez a bizonyítás eltér Fiacco és McCormick bizonyításától.

Az extrapolációs technika kiterjesztésének az volt a célja, hogy egy szukcessziv extrapolációs eljárást dolgozhassunk ki. Ezt a következőképpen értjük. Első lépésben meghatározzunk egy $x^{(0)} = x(r)$ pontot, ahol r értéke nem túl kicsi.

Ezután az

$$/3.27/ \quad x\left(\frac{r}{2}\right) \approx x(r) - \frac{r}{2} h^{oo}(r) = x^{(1)}$$

közelítést alkalmazzuk. Az $x^{(1)}$ ponton át egy $x(r, \theta)$ görbét fektetve meghatározzuk az $x^{(2)}$ pontot. Általában az

$$/3.28/ \quad x^{(k+1)} \approx x^{(k)} - 2^{-1} r^{(k)} h^{oo(k)}$$

iterációt alkalmazzuk, ahol

$$h^{oo(k)} = dx(r^{(k)}, \theta^{(k)}) / dr$$

és $r^{(k)}$ -t ill. $\theta^{(k)}$ -t a 2. szakaszban leirt módon /ld. (2.5) / határozzuk meg. Természetesen konkrét feladat esetén a feladatmegoldó hatékonyabb iterációt is kikísérletezhet más lépéshossz alkalmazására!

Az algoritmusnak az itt megadott formája kisdimenziós tesztfeladatok esetén hatékonyan működött. Az algoritmus matematikai szempontból is megnyugtató tulajdonságokkal rendelkezik. Megmutatható, hogy a $\theta^{(k)}$ sorozat korlátos és $x^{(k)}$ pedig lineárisan konvergál x^* -hoz.

A /3.5/ egyenlet együttható mátrixának szingularitásából eredő nehézségek nem lépnek fel, ha a megoldásban n feltétel aktiv. Legyen pl. $I(\mathbf{x}^*) = \{1, \dots, n\}$.

Az \mathbf{x}^* -ra tett feltételeink mellett következik, hogy ekkor \mathbf{x}^* az R megengedett tartomány csucsa.

$$A \quad \Gamma(\mathbf{x}) = \sum_{i=1}^n u_i^2 \nabla g_i(\mathbf{x}) \nabla g_i'(\mathbf{x})$$

mátrix nonszinguláris, hacsak \mathbf{x} elég közel van \mathbf{x}^* -hoz. Ez következik a $\nabla g_i(\mathbf{x}^*)$ vektorok lineáris függetlenségéből. Ha a \mathbf{h} vektort a /3.5/ egyenlet helyett a

$$\Gamma(\mathbf{x})\mathbf{h} = \nabla f$$

egyenletből határozzuk meg, akkor a megoldás hibája r nagyságrendű. Az extrapolációs technika az adott esetben a

$$g_i(\mathbf{x}^*) = 0 \quad i = 1, \dots, n$$

nemlineáris egyenletrendszer megoldását adja. Érdekes volna megvizsgálni, hogy az extrapolációs technika hatékonyabb-e más egyenletmegoldó módszereknél.

4. Az extrapolációs technika további büntetőfüggvények esetén

Az extrapolációs technikát a reciprok büntetőfüggvény esetén is alkalmazhatjuk, kissé módosított formában. Az idevonatkozó eredményeket részletesebben ismertetjük, mivel ezek Fiacco és McCormick könyvében nem szerepelnek.

A büntetőfüggvényt jelölje

$$/4.1/ \quad I(\mathbf{x}) = \sum_{i=1}^m \frac{1}{g_i(\mathbf{x})} \quad \mathbf{x} \in R^0$$

A segédfüggvény ekkor

$$/4.2/ \quad P(x,r) = f(x) + rI(x).$$

Igaz a következő

17. Tétel. Ha teljesülnek a 13. tétel feltételei, akkor létezik a $P(x,r)$ függvény izolált lokális minimumainak egy $x(r)$ folytonos trajektóriája. Az $x(r)$ görbe $r^{\frac{1}{2}}$ szerint differenciálható $r_0 \leq r \leq 0$ esetén.

Bizonyítás: A tétel első állítása hasonlóan bizonyítható, mint a 13. Tétel megfelelő állítása. Térjünk rá a tétel második részére. Az $x(r)$ trajektória mentén fennáll a

$$/4.3/ \quad \nabla f(x(r)) + r \nabla I(x(r)) = 0$$

egyenlet, ami részletesebben

$$/4.4/ \quad \nabla f(x(r)) - r \sum_{i=1}^m \frac{\nabla g_i(x(r))}{g_i^2(x(r))} = 0$$

Vezessük be az

$$/4.5/ \quad u_i(r) = \frac{r}{g_i^2(x(r))} \quad i = 1, \dots, m$$

jelöléseket. A 6. tételhez hasonlóan látható, hogy $r \rightarrow 0$ esetén $u_i(r)$ tart u_i^* -hoz, ahol u_i^* -gal az x^* ponthoz tartozó Lagrange-szorzókat jelöljük.

Az $x(r)$ görbét differenciáljuk először r szerint. A /4.3/ egyenlet r szerinti differenciálásából a következő közelítő értéket kapjuk:

$$/4.6/ \quad \left(\nabla^2 f(x(r)) + r \nabla^2 I(x(r)) \right) \frac{dx}{dr} = - \nabla I(x(r)) = r^{-1} \nabla f(x(r)).$$

A $P(x, r)$ függvénynek az $x(r)$ pontban vett Hesse-mátrixát jelölje $D(r)$ tehát

$$/4.7/ \quad D(r) = \nabla^2 f(x(r)) + r \nabla^2 I(x(r)).$$

A /4.6/ egyenlet tehát így is írható:

$$/4.8/ \quad D(r) \frac{dx(r)}{dr} = r^{-1} \nabla f(x(r)).$$

A $D(r)$ mátrix invertálható, ezt logaritmikus büntetőfüggvény esetén a 12. tétel állította. A 12. tétel reciprok büntetőfüggvény esetére is általánosítható.

Az $x(r)$ -nek az $s = r^{\frac{1}{2}}$ változó szerinti deriváltját a

$$/4.9/ \quad \frac{dx(r)}{ds} = \frac{dx(r)}{dr} \cdot \frac{dr}{ds}$$

szabály szerint számíthatjuk ki.

Mivel $\frac{dr}{ds} = 2r^{\frac{1}{2}}$, /4.8/ -ből kapjuk a

$$/4.10/ \quad 2(r^{\frac{1}{2}} D(r)) \frac{dx(r)}{ds} = \nabla f(x(r))$$

egyenletet.

Vizsgáljuk meg részletesebben a $D(r)$ mátrixot:

$$/4.11/ \quad D(r) = \nabla^2 f(x(r)) - \sum_{i=1}^m \frac{r}{g_i^2(x(r))} \nabla^2 g_i(x(r)) + \\ + \sum_{i=1}^m \frac{2r}{g_i^3(x(r))} \nabla g_i(x(r)) \nabla g_i'(x(r)).$$

Az $u_i(r)$ értéket /4.5/-ből behelyettesítve és bevezetve a

$$/4.13/ \quad \Gamma(x(r)) = \sum_{i=1}^m u_i(r)^{\frac{3}{2}} \nabla g_i(x(r)) \nabla g_i'(x(r))$$

jelölést /4.12/ így egyszerűsödik:

$$/4.14/ \quad r^{\frac{1}{2}} D(x(r)) = r^{\frac{1}{2}} \nabla^2 L(x(r), u(r)) + \Gamma(x(r)).$$

A /4.10/ egyenletet ugyanúgy elemezhetjük, mint a /3.8/ egyenletet. Így megmutatható, hogy az $x(r)$ -nek az s szerinti deriváltja folytonos és létezik a határértéke $s \rightarrow 0$ esetén. Ebből pedig az $s = 0$ helyen való differenciálhatóság is következik.

Az $x(r)$ görbére tehát a következő közelítés érvényes

$$/4.15/ \quad x(r) \approx x^* + r^{\frac{1}{2}} h_0.$$

A közelítés hibája r nagyságrendű. /4.15/ alapján x^* -ra a következő közelítő értéket kapjuk:

$$/4.16/ \quad x^* \approx \frac{r_1^{\frac{1}{2}} x(r_2) - r_2^{\frac{1}{2}} x(r_1)}{r_1^{\frac{1}{2}} - r_2^{\frac{1}{2}}} \quad (r_1 > r_2 > 0).$$

Az előzőekben az extrapolációs technikát csak az (A) feladattal kapcsolatban dolgoztuk ki. A (B) feladattal kapcsolatban csupán a többé-kevésbé ismert eredmények összefoglalására szorítkozunk. Első tételünk egy vegyes büntetőfüggvénnyel kapcsolatos, amely a következő:

$$/4.17/ \quad P(x,r) = f(x) - r \sum_{i=1}^m \ln g_i(x) + r^{-1} \sum_{j=1}^p h_j^2(x).$$

18. Tétel. Legyenek az f, g_i, h_j függvények kétszer folytonosan differenciálhatók. Az x^* pontban a $\nabla g_i(x^*)$, $i \in I(x^*)$ és $\nabla h_j(x^*)$ vektorok legyenek lineárisan függetlenek. Feltesszük továbbá, hogy az $u_i^* g_i(x^*) = 0$, $i=1, \dots, m$, komplementaritási feltételek szigoruan teljesülnek. Végül feltesszük, hogy x^* -ban teljesül a másodrendű elégséges feltétel. Ekkor létezik a /4.7/ képlettel definiált $P(x,r)$ függvény izolált lokális minimumainak egy folytonos trajektóriája, $x(r)$ amely tart x^* -hoz, amikor $r \rightarrow 0$. Itt r értéke 0 és valamilyen r_0 között változik. Az $x(r)$ trajektória folytonosan differenciálható $r_0 \geq r \geq 0$ esetén.

Vegyük észre, hogy az egyenlőségfeltételekre nem követeltük meg a szigorú komplementaritási feltétel teljesülését. A tétel első felének a bizonyítása megtalálható Fiacco és McCormick könyvében, az $x(r)$ trajektória differenciálhatósága az ott leírt gondolatmenetből könnyen adódik.

Tiszta négyzetes büntetőfüggvény alkalmazása esetén a segédfüggvény

$$/4.18/ \quad P(x,r) = f(x) + r^{-1} \sum_{i=1}^m \left\{ \min[0, g_i(x)] \right\}^2 + r^{-1} \sum_{j=1}^p h_j^2(x)$$

Igaz a következő

19. Tétel. A 18. tétel feltételei mellett létezik a /4.18/ képlettel definiált $P(x,r)$ függvény izolált lokális minimumainak egy folytonos trajektóriája $x(r)$ ($0 < r \leq r_0$) amely tart x^* -hoz, amikor r tart 0-hoz. Az $x(r)$ trajektória folytonosan differenciálható $r_0 \geq r \geq 0$ esetén.

Ez a tétel Fiacco és McCormick könyvében abban formában szerepel, amikor egyenlőségfeltételek nincsenek. Az ott közölt bizonyítás nehézség nélkül átvihető a 19. tétel bizonyítására.

Végül tekintsük a következő vegyes büntetőfüggvényt:

$$/4.19/ \quad P(x, r) = f(x) + r \sum_{i=1}^m \frac{1}{g_i(x)} + r^{-\frac{1}{2}} \sum_{j=1}^p h_j^2(x).$$

A 17. Tétel és a 19. Tétel eredményeit összevetve most már legalábbis szemléletesen világos a két büntetőfüggvény előtt álló paraméterek fenti egyeztetése. Valóban igaz a következő 20. Tétel. A 18. tétel feltételei mellett létezik az /4.19/ képlettel definiált $P(x, r)$ függvény izolált lokális minimumainak egy folytonos trajektóriája $x(r)$, amely tart x^* -hoz, ha $r \rightarrow 0$. Az $x(r)$ görbe $r^{\frac{1}{2}}$ szerint folytonosan differenciálható $r_0 \geq r \geq 0$ mellett.

A fent ismertetett módszer segítségével érdekes eredményt kapunk a $\nabla^2 P(x, r)$ mátrix aszimptotikus viselkedésére. Legyen x^* az (A) feladat egy megoldása és $P(x, r)$ legyen a logaritmikus büntetőfüggvény. Az egyszerűség kedvéért legyenek $1, \dots, k$ az aktív indexek x^* -ben. Legyen M egy $n \times (n-k)$ méretű mátrix, melynek oszlopvektorai kifeszítik a $\nabla g_1(x^*), \dots, \nabla g_k(x^*)$ vektorok által kifeszített altér kiegészítő alterét. Igaz a következő

21. Tétel. A tétel feltételei mellett $(\nabla^2 P(x, r))^{-1}$ limesze $r \rightarrow 0$ esetén

$$M(M' \nabla^2 L(x^*, u^*) M)^{-1} M'.$$

Könnyű belátni, hogy a kapott kifejezés M speciális választásától nem függ. Hasonló eredményt közöl Fletcher a reciprok büntetőfüggvény esetére - bizonyítás nélkül.

5. Érzékenységi vizsgálatok

Fontos kérdés annak a megvizsgálása, hogyan függ egy nemlineáris programozási feladat megoldása a célfüggvény, illetve a feltételi függvények kiértékelésében elkövetett hibáktól. Kicsiny hibahatárok között érzékenységi vizsgálatokkal kapunk választ. Legyen a feladat a következő alakú:

$$/5.1/ \quad \min f(x, \epsilon)$$

feltéve, hogy

$$/5.2/ \quad g_i(x, \epsilon) \geq 0 \quad i = 1, \dots, m,$$

$$/5.3/ \quad h_j(x, \epsilon) \geq 0 \quad j = 1, \dots, p.$$

Érzékenységi vizsgálatokhoz elegendő volna azt az esetet tekinteni, amikor ϵ additív módon szerepel:

$$f(x, \epsilon) = f(x) + \epsilon \quad \text{stb.}$$

De a most következő eredménynek más alkalmazásai is lesznek.

Igaz a következő tétel:

22. Tétel. Legyenek az $f(x, \epsilon)$, $g_i(x, \epsilon)$, $h_j(x, \epsilon)$ $i = 1, \dots, m$, $j = 1, \dots, p$ függvények az x, ϵ változók együttesében folytonosan differenciálhatók. Az $\epsilon = 0$ esetben teljesüljenek a 18. tétel feltételei egy x^* pontban. Ekkor elegendő kis ϵ esetén létezik az /5.1/ /5.2/ /5.3/ feladatoknak egy $x^*(\epsilon)$ szigorú lokális megoldása, úgy hogy $x^* = x^*(0)$ továbbá az $x^*(\epsilon)$ görbe folytonosan differenciálható. /ld [1] /

A tétel egyszerűen bizonyítható az implicit függvényekre vonatkozó tételek alapján, amit a szükséges feltételekre kell alkalmazni. Az érzékenységi vizsgálatban a $\frac{dx^*(\epsilon)}{d\epsilon}$ derivált játszik szerepet.

Ezt is megkaphatjuk a szükséges feltételek differenciálásával, de jó közelítő értéket kapunk a SUMT módszer segítségével. A közelítés alapja Fiacco következő tétele:

23. Tétel. Az /5.1/ /5.2/ /5.3/ feladattal kapcsolatban teljesüljenek a 22. tétel feltételei. Vezessük be a

$$/5.4/ \quad P(x, \varepsilon, r) = f(x, \varepsilon) - r \sum_{i=1}^m \ln g_i(x, \varepsilon) + r^{-1} \sum_{j=1}^p h_j^2(x, \varepsilon)$$

vegyes büntetőfüggvényt, Ekkor az x^* pont egy elegendő kis környezetében létezik a $P(x, \varepsilon, r)$ függvénynek egy egyértelműen meghatározott $x^*(\varepsilon, r)$ szigorú lokális minimuma, ha csak ε, r elég kicsiny. Rögzített r mellett az $x^*(\varepsilon, r)$ görbeiv folytonosan differenciálható és

$$/5.5/ \quad \frac{dx^*(\varepsilon, r)}{d\varepsilon} \rightarrow \frac{dx^*(\varepsilon)}{d\varepsilon} \quad \text{ha } r \rightarrow 0,$$

ε -ban egyenletesen. /ld. [1] /

A tételt az eddigi ismerttetett eszközök alapján könnyű bizonyítanunk. Bizonyításunk különbözik a Fiacco által adott bizonyítástól. Az egyszerűség kedvéért tegyük fel, hogy csak az egyenlőtlenség feltételek szerepelhetnek és ε skalár. $P(x, \varepsilon, r)$ definíciója tehát

$$/5.6/ \quad P(x, \varepsilon, r) = f(x, \varepsilon) - r \sum_{i=1}^m \ln g_i(x, \varepsilon)$$

Rögzített ε esetén $x^*(\varepsilon, r)$ $r \geq 0$ esetén folytonosan differenciálható és az előző eredmények alapján könnyű belátni, hogy $dx^*(\varepsilon, r)/dr$ ε -ban egyenletesen korlátos. Ezért $x^*(\varepsilon, r) \rightarrow x^*(\varepsilon)$ r -ben egyenletes. A 23. tétel bizonyításához elegendő megmutatnunk, hogy $dx^*(\varepsilon, r)/d\varepsilon$ határértéke létezik $r \rightarrow 0$ esetén, mégpedig ε -ban egyenletesen.

Az $x(\varepsilon, r)$ pontot a

$$\nabla_x P(x, \varepsilon, r) = 0$$

egyenletből határozzuk meg. Ezt ε szerint differenciálva a

$$/5.7/ \quad \nabla_{xx}^2 P \cdot \frac{dx^x(\varepsilon, r)}{d\varepsilon} + \nabla_{xr}^2 P = 0$$

egyenletet kapjuk. Ezt a lineáris egyenletet kell $dx/d\varepsilon$ -ra megoldani.

A $\nabla_{xx}^2 P$ mátrixot a /3.7/ alakban akarjuk közelítőleg előállítani. Vezessük be ezért a következő jelöléseket:

$$/5.8/ \quad u_i = u_i(\varepsilon, r) = \frac{r}{g_i}(x(\varepsilon, r)) \quad i = 1, \dots, m,$$

$$/5.9/ \quad A = A(\varepsilon, r) = \nabla_{xx}^2 L(x(\varepsilon, r), u(\varepsilon, r))$$

$$/5.10/ \quad N = N(\varepsilon, r) = (u_1(\varepsilon, r) \nabla g_1(x(\varepsilon, r))), \dots, \\ \dots, u_k(\varepsilon, r) \nabla g_k(\varepsilon, r)$$

ahol $I^* = \{1, \dots, k\}$ az aktív indexek halmazát jelöli.

/Ez x^* egy kicsiny környezetében nyilván független ε -től, mivel szigorú komplementaritási feltétel teljesül/. Ekkor

$$/5.11/ \quad \nabla_{xx}^2 P(x, (\varepsilon, r), \varepsilon, r) \approx A + r^{-1} N N'$$

Az egyenlőség közelítőleg teljesül, ugyanis elhagytuk a $-(r/g_i^2) \nabla g_i \nabla g_i'$ alakú tagokat $i \notin I^*$ -ra. A közelítés hibája legfeljebb $C_1 \cdot r$, ahol C_1 ε -től független pozitív szám.

A $\nabla_{x\varepsilon}^2 P$ vektorra kapjuk, hogy

$$\begin{aligned} /5.12/ \quad \nabla_{x\varepsilon}^2 P &\approx \nabla_{x\varepsilon}^2 L(x,u) + \sum_{i=1}^m \frac{r}{g_i^2} \nabla_x g_i \cdot g_{i\varepsilon} \approx \\ &\approx \nabla_{x\varepsilon}^2 L(x,u) + r^{-1} N g_\varepsilon \end{aligned}$$

ahol

$$g_\varepsilon = \begin{pmatrix} g_{1\varepsilon} \\ \vdots \\ g_{k\varepsilon} \end{pmatrix}$$

A közelítés hibája nyilván legfeljebb $C_2 r$, ahol C_2 egy ε -től független pozitív szám.

Alkalmazzunk az N mátrixra egy Householder-triangularizációt:

$$/5.13/ \quad Q(\varepsilon, r) N(\varepsilon, r) = \begin{pmatrix} R(\varepsilon, r) \\ 0 \end{pmatrix}.$$

Az /5.7/ egyenletet szorozzuk meg balról rQ -val és vezessük be az

$$/5.14/ \quad l = Q \frac{dx(\varepsilon, r)}{d\varepsilon}$$

változót. Így kapjuk az

$$/5.15/ \quad (rQAQ' + QNN'Q') l = -rQ \nabla_{x\varepsilon}^2 L(x,u) - QNg_\varepsilon$$

közelítő egyenletet. Az együtthatómátrix, illetve a jobb-
oldal hibája r^2 nagyságrendű. A 3. szakasz gondolatmenetét szószerint követve kapjuk, hogy az /5.15/ egyenlet l megoldásának s így $dx(\varepsilon, r)/d\varepsilon$ -nak van határértéke, amint $r \rightarrow 0$, sőt ez a határérték ε -ban egyenletesen létezik.

Fiacco megmutatta azt is, hogy

$$/5.16/ \quad \frac{d}{d\varepsilon} u_i(\varepsilon, r) \rightarrow \frac{du_i^*(\varepsilon)}{d\varepsilon} \quad i = 1, \dots, m -re$$

ε -ban egyenletesen. Itt $u_i^*(\varepsilon)$ az optimális $x^*(\varepsilon)$ megoldáshoz tartozó multiplikátorokat jelöli.

III. A MULTIPLIKÁTOR MÓDSZER.

1. Elméleti összefoglalás.

A következőkben a

$$(c) \quad \begin{aligned} & \min f(x) \\ & h_j(x) = 0 \quad j = 1, \dots, p. \end{aligned}$$

feladattal fogunk foglalkozni. Feltesszük, hogy f, h_j kétszer folytonosan differenciálható függvények. A feladat egy izolált lokális megoldását jelölje x^* .

A négyzetes büntetőfüggvény alkalmazása esetén rosszul kondicionált feladatok egy sorozatát kell megoldanunk. A rosszul kondicionáltság az r paraméter csökkentésénél jelentkezik. Hestenes megmutatta, hogy r csökkentése elkerülhető, ha egy más segédfüggvényt építünk fel, az ún. bővített Lagrange függvényt. Hestenes módszerét szokás multiplikátor módszernek is nevezni. Ugyanezt a módszert javasolta Powell, kissé más felfogásban. A multiplikátor módszer azóta is intenzív vizsgálat tárgya.

A feladat vizsgálatának egy természetes utja az optimalitás elsőrendű szükséges feltételéből indul ki. Ennek alakja a (C) feladatra a következő:

$$/1.1/ \quad \nabla f(x^*) + \sum_{j=1}^p w_j^* \nabla h_j(x^*) = 0$$

$$h_j(x^*) = 0 \quad j = 1, \dots, p.$$

Itt w_j^* -gal a Lagrange - szorzókat jelöltük. Az /1.1/ feltételrendszer $n+p$ nemlineáris egyenletet tartalmaz és ugyanennyi az ismeretlenek száma is. Nemlineáris egyenletrendszerek megoldására ismerünk jól kidolgozott módszereket. Ezek alkalmazhatóságának egyik alapfeltétele, hogy az egyenletrendszer Jacobi-mátrixa az (x^*, w^*) pontban ne legyen szinguláris. Ez elméletileg is érdekes kérdés, ezért a következő tételben erre adunk egy elégséges feltételt.

24. Tétel. Legyen x^* a (C) feladat egy izolált lokális megoldása. A $\nabla h_j(x^*)$ vektorok legyenek lineárisan függetlenek és teljesüljön x^* -ban az optimalitás másodrendű elégséges feltétele. Ekkor léteznek olyan w_j^* Lagrange-szorzók, hogy teljesül az /1.1/ feltételrendszer és az /1.1/ egyenletrendszer Jacobi-mátrixa az (x^*, w^*) pontban nem-szinguláris.

Bizonyítás: A 3. tétel alapján tudjuk, hogy léteznek olyan w_j^* Lagrange-szorzók, amelyekkel teljesül az /1.1/ feltételrendszer. A (C) feladat Lagrange-függvényét jelölje $L(x, w)$ azaz

$$/1.2/ \quad L(x, w) = f(x) + \sum_{j=1}^p w_j h_j(x)$$

A $h_j(x^*)$ vektorokból alkotott mátrixot jelölje $H(x^*)$, azaz

$$/1.3/ \quad H(x^*) = (\nabla h_1(x^*), \dots, \nabla h_p(x^*)).$$

Az /1.1/ egyenletrendszer Jacobi-mátrixa a következő alakú:

$$G = \begin{pmatrix} \nabla_{xx}^2 L(x^*, w^*) & H(x^*) \\ H'(x^*) & 0 \end{pmatrix}$$

A $C = \nabla_{xx}^2 L(x^*, w^*)$ és az $A = H(x^*)$ mátrixokra teljesülnek a 8. tétel feltételei, ezért a 8. tétel szerint G nem-szinguláris, amit bizonyítani akartunk.

A most bizonyított tétel alapján megpróbálhatjuk az /1.1/ egyenletrendszert a Newton-módszerrel megoldani. Ehhez azonban egy $(n+p) \times (n+p)$ méretű mátrixot kell invertálni, ami nehézkes lehet. Ez az egyik oka annak, hogy számos speciális módszert fejlesztettek ki a (C) feladat ill, az azzal ekvivalens /1.1/ egyenletrendszer megoldására.

Definiáljuk a bővített Lagrange-függvényt a

$$/1.5/ \quad Q(x, w, k) = f(x) + \sum_{j=1}^p w_j h_j(x) + k \sum_{j=1}^p h_j^2(x)$$

képlettel. Itt w egy p dimenziós vektor, amelynek a komponensei w_1, \dots, w_p , a k pedig egy pozitív szám. Igaz a következő tétel.

25. Tétel. Teljesüljenek a 24. tétel feltételei.

Ekkor

$$/1.6/ \quad \nabla_x Q(x^*, w^*, k) = 0$$

és a

$$\nabla_{xx}^2 Q(x^*, w^*, k)$$

mátrix pozitív definit, hacsak k elég nagy.

Bizonyítás: A 24. tétel alapján léteznek olyan Lagrange-szorzók, amelyekkel fennáll a

$$\nabla_x L(x^*, w^*) = 0$$

feltétel. Innen kapjuk, hogy

$$\nabla_x Q(x^*, w^*, k) = \nabla_x L(x^*, w^*) + 2k \sum_{j=1}^p h_j(x^*) \nabla_x h_j(x^*) = 0$$

mivel $h_j(x^*) = 0$, $j = 1, \dots, p$, tehát teljesül az /1.6/ feltétel. Még egyszer differenciálva x szerint és figyelembe véve, hogy $h_j(x^*) = 0$ kapjuk a

$$/1.7/ \quad \nabla_{xx}^2 Q(x^*, w^*, k) = \nabla_{xx}^2 L(x^*, w^*) + 2k \sum_{j=1}^p \nabla_x h_j(x^*) \nabla_x h_j(x^*)$$

összefüggést.

A jobboldalon álló kifejezésre alkalmazni fogjuk a Finsler-lemmát. Tekintsük a

$$/1.8/ \quad \nabla_x h_j(x^*) \cdot v = 0 \quad j = 1, \dots, p$$

feltételek által meghatározott alteret. A

$$/1.9/ \quad C = \nabla_{xx}^2 L(x^*, w^*)$$

mátrix pozitív definit az /1.8/ feltételekkel meghatározott altéren, mivel teljesül az optimalitás másodrendű elégséges feltétele. Megmutatjuk, hogy a

$$/1.10/ \quad D = \sum_{j=1}^p \nabla_x h_j(x^*) \nabla_x' h_j(x^*)$$

mátrixra teljesülnek a Finsler-lemmában megkivánt feltételek. Világos, hogy ha v kielégíti az /1.8/ feltételeket, akkor $Dv = 0$. Másrészt legyen y az /1.8/ feltételek által meghatározott altér kiegészítő alterének egy nem-nulla vektora:

$$/1.11/ \quad y = \sum_{j=1}^p \lambda_j \nabla_x h_j(x^*) \neq 0$$

Meg kell mutatnunk, hogy $y'Dy$ pozitív. Tegyük fel az ellenkezőjét. Ekkor az

$$/1.12/ \quad y'Dy = \sum_{j=1}^p (y' \nabla_x h_j(x^*))^2 = 0$$

egyenlőségből kapjuk, hogy

$$/1.13/ \quad y' \nabla_x h_j(x^*) = 0 \quad j = 1, \dots, p \text{-re.}$$

Beszorozva a j -edik egyenlet λ_j -vel és összeadva $j = 1, \dots, p$ -re az

$$/1.14/ \quad y'y = 0$$

eredményt kapjuk, ami ellentmondás. Ezzel beláttuk, hogy a Finsler-lemma valamennyi feltétele teljesül. A Finsler-lemma állítása szerint az $F = C + kD$ mátrix pozitív definit, ha k elég nagy. Ez pedig éppen a bizonyítandó állítás. Ezzel a 25. tételt bebizonyítottuk.

A tételt a következőképpen értékelhetjük. A 3. tétel szerint léteznek olyan w_j^* ($j = 1, \dots, p$) Lagrange-szorozók, hogy (x^*, w^*) az $L(x, w)$ függvény stacionárius pontja. Ennek a minőségét azonban nem ismerjük. Ezzel szemben a $Q(x, w, k)$ függvény esetén tudjuk, hogy az (x^*, w^*) pontban Q -nak az x változó szerint lokális minimuma van.

2. A multiplikatormódszer levezetése.

A multiplikatormódszer alapgondolata az, hogy ha w^* ismert lenne, akkor x^* -ot egyetlen feltétel nélküli minimalizálással meg tudjuk határozni. Sajnos w^* közvetlenül nem ismert, ezért azt iteratív uton kell meghatározni.

Ha w a w^* multiplikatőr egy elég jó közelítése, akkor a

$$/2.1/ \quad \nabla_x Q(x, w, k) = 0$$

egyenletnek x^* egy kis környezetében egyetlen $x(w)$ megoldása van. Ez következik az implicit függvényekre vonatkozó tételekből, mivel $\nabla_{xx}^2 Q(x^*, w^*, k)$ nem szinguláris. Az f, h_j függvények kétszer folytonosan differenciálhatók, ezért a $\nabla_{xx}^2 Q(x(w), w, k)$ mátrix is pozitív definit, hacsak w elég közel van w^* -hoz. Ezért $x(w)$ megoldása a

$$/2.2/ \quad \min_x Q(x, w, k)$$

feltétel nélküli minimalizálási problémának is. Ez az észrevétel döntő abból a szempontból, hogy $x(w)$ -t csupán az f, h_j függvényértékek alapján kiszámíthatjuk.

A w^* vektort az jellemzi, hogy a

$$/2.3/ \quad h_j(x(w)) = 0$$

egyenlőség teljesül. Valóban, ha teljesül az /2.3/ egyenlőség, akkor $x^* = x(w^*)$ megengedett pont, továbbá kielégíti a Lagrange-féle szükséges feltételt is:

$$\nabla_x L(x^*, w^*) = \nabla_x Q(x^*, w^*) = 0.$$

A w^* vektor meghatározásának feladatát a /2.3/ nemlineáris egyenletrendszer megoldására vezettük vissza. Sajnos,

ebben szerepel egy $x(w)$ implicit módon definiált függvény, ezért a Newton-módszer helyett annak egy közelítését fogjuk kidolgozni.

Mielőtt erre rátérnénk, a w^* vektor egy további jellemzését fogjuk adni. Vezessük be a

$$/2.4/ \quad \Psi(w) = Q(x(w), w, k)$$

függvényt. Tekintjük a

$$/2.5/ \quad \max_w \Psi(w)$$

feltétel nélküli maximalizálási feladatot.

Rockafellartól származik a következő tétel /ld. [32] /:

26. Tétel. A 24. Tétel feltételeinek teljesülése mellett w^* a /3.5/ probléma egy izolált lokális megoldása.

Ennek az észrevételnek az a jelentősége, hogy a w^* meghatározására alkalmazhatók a korszerű függvényminimalizálási módszerek.

A 26. Tétel bizonyításához számítsuk ki az $x(w)$ függvény parciális deriváltjait. A

$$\nabla_x Q(x(w), w, k) = 0$$

egyenlet w szerinti differenciálásával a

$$\nabla_{xx}^2 Q \frac{\delta x}{\delta w} + \nabla_{xw}^2 Q = 0$$

egyenletet kapjuk. A

$$/2.6/ \quad H = (\nabla_x h, \dots, \nabla_x h_p)$$

jelölés felhasználásával kapjuk, hogy

$$/2.7/ \quad \frac{\delta x}{\delta w} = -(\nabla_{xx}^2 Q)^{-1} H$$

A jobboldalt értelemszerűen a $w, x(w)$ pontban kell kiértékelni.

Számítsuk most ki a $\Psi(w)$ függvény deriváltjait a $w=w^*$ pontban:

$$/2.8/ \quad \nabla_w \Psi = \nabla_x Q \frac{\delta x}{\delta w} + \nabla_w Q = 0 ,$$

mivel

$$\nabla_x Q = 0 \quad \text{és} \quad \nabla_w Q = h(x(w^*)) = 0. \quad A$$

második deriváltakra pedig a

$$/2.9/ \quad \nabla_{ww}^2 \Psi = \left(\frac{\delta x}{\delta w} \right)' \left(\nabla_{xx}^2 Q \right) \left(\frac{\delta x}{\delta w} \right) + 2 \nabla_{xw}^2 Q \frac{\delta x}{\delta w}$$

kifejezést kapjuk. A deriválásból adódó egyéb tagok 0-val egyenlők.

A /2.7/ összefüggés felhasználásával azt kapjuk, hogy

$$/2.10/ \quad \nabla_{ww}^2 \Psi = - H' \left(\nabla_{xx}^2 Q \right)^{-1} H ,$$

ez a mátrix pedig negatív definit.

Visszatérve a

$$/2.11/ \quad h(x(w)) = 0$$

egyenlet numerikus megoldására megmutatjuk, hogy a Newton-módszernek igen jó approximációja a

$$/2.12/ \quad \delta w = 2kh(x(w))$$

iterációval definiált módszer. Ezen az iteráción alapuló megoldást nevezzük multiplikátormódszernek.

A /2.12/ iteráció levezetéséhez alkalmazzuk előbb a Newton-módszert a /2.11/ egyenlet megoldására. A baloldali Jacobi-mátrixára a

$$/2.13/ \quad G = - H' (\nabla_{xx}^2 Q)^{-1} H$$

mátrixot kapjuk.

A /2.13/ egyenlőség jobboldalán a

$$/2.14/ \quad (2k)^{-1} \nabla_{xx}^2 Q = \varepsilon + HH' .$$

átalakítást végezzük. Itt az

$$/2.15/ \quad \varepsilon = (2k)^{-1} \nabla_{xx}^2 L$$

mátrix elemei $O(k^{-1})$ nagyságrendűek. Így a

$$/2.16/ \quad 2kG = - H' (\varepsilon + HH')^{-1} H$$

összefüggést kapjuk.

A H mátrixra alkalmazzunk egy Householder triangularizációt:

$$/2.17/ \quad Q H = \begin{pmatrix} R \\ 0 \end{pmatrix}$$

ahol R felsőháromszög alakú nonszinguláris mátrix, Q pedig ortogonális.

Igy a

$$/2.18/ \quad 2kG = - R' (\varepsilon' + RR')^{-1} R = - I + O(k^{-1})$$

összefüggés adódik, ahol I az egységmátrix. Így kapjuk a következő, Powell-től származó tételt:

27. Tétel. A 24. Tétel feltételei mellett ha w elég közel van w^* -hoz és $k > k_0$, akkor

$$/2.19/ \quad | - G_{st} - \delta_{st} / 2k | < ck^{-2}$$

ahol c konstans, δ_{st} pedig a Kronecker-szimbólumot jelöli.

/2.19/ invertálásával

$$(2kG)^{-1} = -I + O(k^{-1})$$

adódik, innen pedig

$$/2.20/ \quad -G^{-1} = 2kI + O(1)$$

adódik. A közelítés hibája tehát nem tart 0-hoz, amint $k \rightarrow \infty$, de a relatív hiba 0-hoz tart. A /2.20/ közelítés levezetésével most már megalapoztuk a /2.12/ multiplikátor-módszert.

Összefoglalva a módszer a következő: kiindulunk a w^* multiplikátor egy w_0 közelítéséből. Az általános k -dik lépésben a már ismert w_k mellett meghatározzuk $x(w_k)$ -t, majd kiszámítjuk w_{k+1} -et a /2.12/ képlet felhasználásával.

A multiplikátormódszer alkalmazásánál kérdéses az első w_0 közelítés megválasztása. Ehhez nyújthat segítséget egy Fletcher által javasolt módszer. A kiindulópontunk az optimalitás elsőrendű szükséges feltétele, amelyet most a

$$/2.21/ \quad \nabla f(x^*) + H(x^*)w^* = 0$$

alakban írunk fel, ahol $H(x^*)$ a /2.11/ képlettel értelmezett $n \times p$ -es mátrix. Szorozzuk be a /2.21/ egyenletet $H'(x^*)$ -gal balról. Kapjuk a

$$H'(x) \nabla f(x^*) + H'(x^*)H(x^*)w^* = 0$$

egyenletet. Feltevésünk szerint $H(x^*)$ rangja p , ezért $H'(x^*)H(x^*)$ nem-szinguláris. Így w^* kifejezhető a következő módon:

$$/2.22/ \quad w^* = -(H'(x^*)H(x^*))^{-1} H'(x^*) \nabla f(x^*) .$$

A jobboldalon álló kifejezés értelmezve van bármely x^* -hoz közeleső x pontra is. Tehát a multiplikátorok egy jó becslését kapjuk a

$$/2.23/ \quad w(x) = -(H'(x)H(x))^{-1} H'(x) \nabla f(x)$$

képlet alapján. Itt $w(x)$ egy p dimenziós vektor, amelynek komponenseit $w_1(x), \dots, w_p(x)$ -szel jelöljük.

Fletcher továbbment okoskodásaiban és bebizonyította a következő tételt:

28. Tétel: A 24. Tétel feltételeinek teljesülése mellett az

$$U(x) = f(x) + \sum_{i=1}^p w_i(x)h_i(x)$$

függvénynek az x^* pont stacionárius pontja. Továbbá ha k elég nagy konstans, akkor x^* a

$$V(x) = f(x) + \sum_{i=1}^p w_i(x)h_i(x) + k \sum_{i=1}^p h_i^2(x)$$

függvénynek szigorú lokális minimumhelye.

3. A bővített Lagrange-függvény további vizsgálata.

A multiplikátormódszert Powell egy más formában vezette le. A bővített Lagrange-függvény helyett ő a

$$/3.1/ \quad Q(x, c, k) = f(x) + k \sum_{j=1}^p (h_j(x) - c_j)^2$$

segédfüggvényt vezette be. A különbség valójában csak formai, a

$$/3.2/ \quad w_j = -2 kc_j$$

választással a /3.1/ függvény bővített Lagrange-függvény alakjában is írható. Mégis a Powell által adott alak általánosításra alkalmasabb.

A multiplikátormódszert egyenlőtlenségfeltételek esetére is általánosítani akarjuk.

Tekintsük az (A) feladatot. A közönséges büntetőfüggvény helyett vegyük a következő segédfüggvényt:

$$/3.3/ \quad Q(x, c, k) = f(x) + k \sum_{i=1}^m (g_i(x) - c_i)^2$$

Itt a_i az a_i szám negatív részét jelenti. Ezt a segédfüggvényt Rockafellar illetve Fletcher vezették be. Az optimum-pontot x^* , a megfelelő Lagrange-szorzókat u_i^* jelöli. A c^* vektort a

$$/3.4/ \quad 2 kc_i^* = u_i^*$$

egyenletekkel definiáljuk. Fletcher bebizonyította a következő tételt. /ld [10] /

29. Tétel. Az (A) feladatra teljesüljenek a 13. Tétel feltételei. Ekkor létezik olyan k_0 szám, hogy $k > k_0$ esetén az x^* megoldás a $Q(x, c^*, k)$ függvény szigorú lokális minimuma.

A 30. és a 31. Tételt nyomdatechnikai okokból mellőzzük.

4. A multiplikátormódszer alkalmazásának numerikus kérdései.

Tegyük fel, hogy az f, h_j függvények második deriváltjai könnyen kiszámíthatók. Ekkor az $x(w)$ meghatározása történhet Newton-módszerrel. Ugyanígy a $\Psi(w)$ függvény maximalizálása is történhet Newton-módszerrel. Számítsuk ki a $\Psi(w)$ függvény deriváltjait:

$$/4.1/ \quad \nabla_w \Psi(w) = \nabla_x Q \frac{\delta x}{\delta w} + \nabla_w Q = \nabla_w Q = h$$

mivel $\nabla_x Q(x(w), w, k) = 0$. Továbbá

$$/4.2/ \quad \nabla_{ww}^2 \Psi = \nabla_x Q \frac{\delta^2 x}{\delta w^2} + \nabla_{xx}^2 Q \frac{\delta x}{\delta w} + \nabla_{ww}^2 Q = \nabla_{xx}^2 Q \frac{\delta x}{\delta w} = - (F^{-1} N_F)$$

mivel $\nabla_{ww}^2 Q = 0$ bármely (x, w) -re. Így a Newton-módszer szerint

$$/4.3/ \quad \delta w = (F^{-1} H_F)^{-1} h .$$

Megjegyezzük, hogy ugyanezt az eredményt kapjuk akkor is, ha a Newton-módszert a /2.3/ nemlineáris egyenletrendszerre alkalmazzuk. Ezeket a képleteket megtaláljuk Fletcher dolgozatában is.

Az $x=x(w)$ felületdarab linearizálásával kapjuk, hogy az új $x(w+\delta w)$ értékeket közelítőleg a

$$/4.4/ \quad \delta x^{(1)} = \frac{\delta x}{\delta w} \delta w = - F^{-1} H (F^{-1} H_F)^{-1} h$$

korrekcióval kell kiszámítani. Megjegyezzük, hogy ha az /1.1/ $n+p$ dimenziós nemlineáris egyenletrendszerre alkalmazzuk a Newton-módszert az $(x(w), w)$ pontban, akkor a /4.3/ /4.4/ korrekciókat kapjuk.

A Függelékben leírt /1.6/ /1.7/ dekompozíciós eljárást fogjuk alkalmazni. Az $x(w)$ vektor meghatározására a Newton-módszert alkalmazzuk, tehát az iterációt a

$$/4.5/ \quad \delta x^{(2)} = - (\nabla_{xx}^2 Q)^{-1} (\nabla_x Q) = - F^{-1} \nabla_x Q$$

képlet definiálja. A jobboldalt itt egy (x, w) pontban értékeljük ki. Az x vektor korrekcióját végül a

$$/4.6/ \quad \delta x = \delta x^{(1)} + \delta x^{(2)}$$

képlettel adjuk meg. A Függelék 36. tétele alapján könnyen igazolható a következő.

32. Tétel. A /4.3/ /4.6/ korrekciós formulákkal definiált iteratív eljárás konvergens.

Most vizsgáljuk meg azt az esetet, amikor csak az f, h_j függvények és első deriváltjaik számíthatók ki. A multiplikátormódszer szerint a w korrekciójára a

$$/4.7/ \quad \delta w = 2kh(x)$$

képletet javasoljuk. Vegyük észere, hogy a jobboldalon x nem szükségképpen azonos $x(w)$ -vel. A δx korrekciót a Függelékben közölt /1.6/ /1.7/ eljárás alapján két részből tesszük össze. Egyrészt a

$$/4.8/ \quad \nabla_x Q(x, w, k) = \text{const}$$

Ljapunov - felülethez megszerkesztjük a δw elmozdulásnak megfelelő érintővektort. Ez nyilván

$$/4.9/ \quad \delta x^{(1)} = - (\nabla_{xx}^2 Q)^{-1} (\nabla_{xw}^2 Q) \delta w = -2kF^{-1} Hh(x).$$

Másrészt a $\nabla_x Q(x(w), w, k) = 0$ egyenlet megoldásához közeledünk, ha a

$$/4.10/ \quad \delta x^{(2)} = - G \nabla_x Q$$

korrekciót alkalmazzuk, ahol G a $(\nabla_{xx}^2 Q)^{-1}$ mátrix egy jó /mindenesetre pozitív definit/ közelítése. Végül definiáljuk a

$$/4.11/ \quad \delta x = \delta x^{(1)} + \delta x^{(2)}$$

korrekciót.

Összefoglalva a

$$\delta w = 2kh(x)$$

$$/4.12/ \quad \delta x = -2kF^{-1}Hh(x) - GQ_x$$

korrekciót kapjuk. A Függelékben közölt 36. Tétel alapján a /4.12/ képlet alapján végzett iteratív eljárás konvergens. Sajnos /4.12/-ben a második deriváltakat tartalmazó F mátrix inverze is szerepel. Ezért a következő módosítást javasoljuk.

$$/4.13/ \quad \begin{aligned} \delta w &= 2kh(x) \\ \delta x &= -2kGHh(x) - GQ_x \end{aligned}$$

Igaz a következő

33. Tétel. A /4.13/ korrekciós képlettel definiált iteratív eljárás konvergens.

Bizonyítás. Jelöljük az /4.12/ illetve /4.13/ képletek jobboldalait $s(z)$ -vel, illetve $t(z)$ -vel. Könnyű látni, hogy $z^x = (x^x, w^x)$ mind a

$$\dot{z} = s(z)$$

mind pedig a

$$\dot{z} = t(z)$$

egyenletnek egyensúlyi pontja. Mivel pedig z^x az első egyenletnek stabil egyensúlyi pontja és fennáll a

$$/4.14/ \quad \|t(z)\| \leq (1 + \varepsilon) \|s(z)\| \quad \varepsilon > 0$$

egyenlőtlenség, ha csak G elegendő jó közelítése F^{-1} -nek, azért z^x /4.13/ -nek is stabil egyensúlyi pontja.

A /4.13/ -iterációs módszer alkalmazásához csak első deriváltakra van szükségünk. A módszer approximációja a Newton-módszernek, amelyet a

$$h(x(w)) = 0$$

egyenletre alkalmazzuk, de a $-G \nabla_x Q$ tag feltehetően javítja a konvergencia tulajdonságokat pl. abban az értelemben, hogy nagyobb a konvergencia-tartomány. Ezek a kérdések azonban még tisztázatlanok.

Ha a feladatban szereplő függvények deriváltjai könnyen kiszámíthatók, akkor a G becslés javítására használhatjuk a Broyden-féle módszereket, vagy a hagyományosabb Davidon-Fletcher-Powell módszert. Az utóbbi módszernek egy deriváltmentes változatát is kidolgozta Stewart, így a multiplikatormódszer olyankor is jól alkalmazható, amikor csupán a feladatban szereplő függvények értékei számíthatók ki könnyen.

IV. Szekvenciális módszerek alkalmazása a sztohasztikus programozásban

1. A sztohasztikus approximáció néhány módszere.

Ebben a fejezetben azt vizsgáljuk meg, hogyan kell a szekvenciális módszereket hatékonyan alkalmazni sztohasztikus programozási feladatok megoldására. A sztohasztikus programozás modelljeinek jó összefoglalását tartalmazza Prékopa András munkája. /ld [28] /

A sztohasztikus programozás feladataiban a feltételi függvények ill. a célfüggvény igen bonyolult integrálkifejezéseket tartalmaznak. Ezeknek Monte-Carlo módszerrel történő integrálása igen időigényes, más numerikus integrálási eljárás pedig a magas dimenziószám miatt nem alkalmazható. A Monte Carlo-módszer és egy nemlineáris programozási módszer egyidejű alkalmazása ezért célszerű.

Igy egy ún. sztohasztikus approximációs eljárást kapunk. A következőkben összefoglaljuk a sztohasztikus approximáció elméletének főbb eredményeit.

Az első klasszikus eredmény a Robbins-Monroe eljárás. Az eljárás célja egy

$$/1.1/ \quad f(x) = 0$$

nemlineáris egyenlet megoldása, ahol $f(x)$ egy regressziós függvény, azaz

$$/1.2/ \quad f(x) = M(Y(x))$$

ahol $Y(x)$ egy realizálható valószínűségi változó, M pedig a zárójelben álló kifejezés várható értékét jelöli. A Robbins-Monroe eljárás esetén x skalárváltozó kell, hogy legyen. Később Blum többdimenziós eljárásokat dolgozott ki. Az egyik legáltalánosabb konstrukció Sz.A. Ivanovtól származik, ezt fogjuk ismertetni.

Először egy determinisztikus eljárásra vonatkozó eredményt mutatunk be.

Legyen x^* egy meghatározandó vektor, amely lehet pl. egy nemlineáris egyenlet gyöke. Legyen adott továbbá egy algoritmus, amelynek alakja

$$/1.3/ \quad x_{n+1} = x_n + \alpha_n r(x_n)$$

Ivankov tétele az /1.3/ típusu iteratív eljárások konvergenciáját vizsgálja. /ld [4] /

34. Tétel. Tegyük fel, hogy az /1.3/ iteratív eljárásra teljesülnek a következő feltételek: Az $r(x)$ vektorfüggvény folytonos $x \in R$ esetén, ahol R egy kompakt halmaz, és $r(x) = 0$ akkor és csak akkor, ha $x = x^*$. Továbbá létezik olyan folytonosan differenciálható nemnegatív konvex $U(x)$ függvény, melyre $U(x) = 0$ akkor és csak akkor, ha $x = x^*$ és

$$/1.4/ \quad r'(x) \nabla_x U < 0 \quad \text{ha} \quad x \neq x^*$$

Végül

$$/1.5/ \quad \alpha_n \rightarrow 0, \quad \sum_{n=1}^{\infty} \alpha_n = \infty$$

és $x_n \in R$ minden n -re. Ekkor az /1.3/ iteratív eljárás konvergens, azaz $x_n \rightarrow x^*$ $n \rightarrow \infty$ esetén.

Megjegyezzük, hogy a tételt Ivankov általánosabb alakban fogalmazza meg: $r(x)$ többértékű függvény is lehet.

Az /1.3/ eljárás egy sztohasztikus változatát a következőképpen szerkesztjük meg: Legyen $\xi_n = \xi(x_n)$ egy valószínűségi változó, melyre

$$/1.6/ \quad M(\xi_n) = r(x_n)$$

és tekintsük az

$$/1.7/ \quad x_{n+1} = x_n + \alpha_n \xi(x_n)$$

eljárást. Erre igaz a következő

35. Tétel. Tegyük fel, hogy teljesülnek a 34. tétel feltételei. Tegyük fel, továbbá, hogy a ζ_n valószínűségi változók függetlenek, szórásaik korlátosak:

$$/1.8/ \quad D^2(\xi_n) < K$$

valamilyen K -val, és

$$/1.9/ \quad \sum \alpha_n^2 < \infty .$$

Ekkor az x_n sorozat egy valószínűséggel tart x^* -hoz.

A 35. tétel alapján fogunk a következő pontban sztohasztikus programozási feladatokat megoldani. Az eljárások alapja a következő lesz: alkalmazzunk egy determinisztikus eljárást, amely a 34. tétel feltételeinek eleget tesz. Az $r(x)$ vektorra Monte-Carlo módszerrel adunk torzítatlan becslést, s ennek felhasználásával alkalmazzuk az /1.7/ sztohasztikus eljárást. Az α_n lépéshossz szokásos választása $\alpha_n = \frac{1}{n}$.

Például az /1.1/ feladat megoldására az

$$x_{n+1} = x_n - \frac{1}{n} f(x_n)$$

determinisztikus eljárás alapján az

$$/1.10/ \quad x_{n+1} = x_n - \frac{1}{n} Y(x_n)$$

sztohasztikus approximációs eljárást szerkeszthetjük. Ez a Robbins-Monroe eljárás.

Ha a determinisztikus feladat függvényminimalizálás /feltétel nélküli vagy feltételes/ úgy az /1.3/ típusú eljárásokban elsőrendű parciális deriváltak ismerete szüksé-

ges. Deriváltmentes eljárást kaphatunk legegyszerűbben úgy, hogy a deriváltakat differenciálhányadossal helyettesítjük. Egy ilyen eljárás sztohasztikus változata a Kiefer-Wolfowitz eljárás, amelyet csak nagy vonalakban foglalunk össze. Tekintsük a

$$/1.11/ \quad \min f(x)$$

feladatot, ahol x skalárváltozó és

$$/1.12/ \quad f(x) = M(Y(x)).$$

Ekkor

$$/1.13/ \quad f'(x) \approx M(Y(x+c) - Y(x-c)) / 2c$$

ezért javasoljuk az

$$/1.14/ \quad x_{n+1} = x_n - \frac{a_n}{c_n} (Y(x+c_n) - Y(x-c_n))$$

eljárást. Az a_n, c_n együtthatók egy célszerű választása

$$/1.15/ \quad a_n = \frac{a}{n} \quad c_n = \frac{c}{n^{1/4}} .$$

A Kiefer-Wolfowitz eljárást 1952-ben publikálták, azóta számos hatékony deriváltmentes algoritmust dolgoztak ki. Érdekes volna ezek sztohasztizálhatóságát megvizsgálni.

2. Veszteségfüggvényes és megbízhatósági jellegű modellek

A veszteségfüggvényes modellek közös jellemzője, hogy a célfüggvény egy determinisztikus függvény és egy valószínűségi változó várható értékének az összege. Az utóbbi függvényt egy integrálkifejezés definiálja, amelynek egy tipikus alakja a következő:

$$/2.1/ \quad \int \dots \int (\beta - Ax) d\Phi(\beta) \\ \beta \geq Ax$$

Itt β valószínűségi változó, $\phi(\beta)$ az eloszlásfüggvénye. Ha a /2.1/ típusú kifejezéseket Monte-Carlo módszerrel értékeljük ki, akkor a célfüggvényre ill. a gradiensére a következő előállítást kapjuk:

$$\begin{aligned} /2.2/ \quad f(\mathbf{x}) &= M(\xi(\mathbf{x})) \\ \nabla_{\mathbf{x}} f(\mathbf{x}) &= M(\zeta(\mathbf{x})) \end{aligned}$$

Itt $\xi(\mathbf{x}), \zeta(\mathbf{x})$ r -től függő valószínűségi változók, amelyek véletlenszámgenerátor segítségével realizálhatók.

A feladatot a

$$/2.3/ \quad \min f(\mathbf{x})$$

$$/2.4/ \quad g_i(\mathbf{x}) \geq 0$$

alakban megfogalmazva azt vizsgáljuk, hogyan alkalmazhatók a SUMT módszer ill. a multiplikátormódszer.

Az /2.3/ /2.4/ feladat megoldására alkalmazzuk pl. a logaritmusos büntetőfüggvény módszert, a segédfüggvényt jelölje $P(\mathbf{x}, r)$. Vezessük be a

$$/2.5/ \quad \hat{P}(\mathbf{x}, r) = \xi(\mathbf{x}) - r \sum_{i=1}^m \ln g_i(\mathbf{x})$$

$$\nabla \hat{P}(\mathbf{x}) = \zeta(\mathbf{x}) - r \sum_{i=1}^m \frac{\nabla g_i(\mathbf{x})}{g_i(\mathbf{x})}$$

véletlen függvényt. Világos, hogy

$$/2.6/ \quad M(\hat{P}(\mathbf{x}, r)) = P(\mathbf{x}, r) \quad \text{és}$$

$$M(\nabla \hat{P}(\mathbf{x})) = \nabla_{\mathbf{x}} P(\mathbf{x}, r).$$

Igy egy Robbins-Monroe típusu eljárást alkalmazhatunk, melynek képlete esetünkben

$$/2.8/ \quad x_{n+1} = x_n - \frac{1}{n} \nabla \hat{P}(x_n).$$

Az x_n sorozat a 35. tétel feltételei mellett egy valószínűséggel konvergál $x^*(r)$ -hez, a $P(x,r)$ függvény minimumához.

Ha már sikerült az aktív indexek halmazát I^* -ot megjelölni, akkor célszerű áttérni a multiplikátormódszerre. Az egyszerűség kedvéért tegyük fel, hogy $I^* = \{1, \dots, p\}$, legyen $h_j(x) = g_j(x)$ $j=1, \dots, p$ -re és tekintsük a

$$/2.9/ \quad \begin{aligned} & \min f(x) \\ & h_j(x) = 0 \quad j = 1, \dots, p \end{aligned}$$

feladatot.

A bővített Lagrange-függvényt jelölje $Q(x,w,k)$, és vezessük még be a

$$/2.10/ \quad \hat{Q}(x,w,k) = f(x) + \sum_{j=1}^p w_j h_j(x) + k \sum_{j=1}^p h_j^2(x)$$

véletlen függvényt. Világos, hogy

$$/2.11/ \quad Q(x,w,k) = M\hat{Q}(x,w,k)$$

ezért a $Q(x,w,k)$ függvény $x(w)$ minimumának a meghatározására alkalmazható a Kiefer-Wolfowitz eljárás. Ha még az $f(x)$ célfüggvény gradiensére is van torzítatlan becslésünk, akkor Robbins-Monroe típusu eljárás alkalmazható. Ez eddig teljesen analóg a büntetőfüggvények módszerével.

A multiplikátormódszer általunk kidolgozott /4.13/ módosítása is sztohasztikus eljárássá fejleszthető.

Definiáljuk ugyanis a

$$/2.12/ \quad \hat{\delta x} = - 2kGHh(x) - GQ_x$$

véletlen irányt. Világos, hogy

$$/2.13/ \quad M(\hat{\delta x}) = \delta x$$

ezért a 35. tételben leirt konstrukció alkalmazható és így egy sztohasztikus eljárást kapunk.

Megbizhatósági jellegű modellek esetén egyes feltételi függvényekben szerepelnek a véletlen hatások. Egy sokat vizsgált példa a

$$/2.14/ \quad g_1(x) = P(Ax \geq \beta) - p$$

feltételi függvény. Itt A egy determinisztikus mátrix, x egy döntési változó, β véletlen vektor, $P(\quad)$ a zárójelben álló esemény valószínűségét jelöli, p pedig egy előirt megbizhatósági szintet. A jobboldal Monte-Carlo módszerrel történő kiértékelése esetén a $g_1(x)$ függvényre egy $\xi(x)$ torzítatlan becslést kapunk:

$$/2.15/ \quad g_1(x) = M(\xi(x)) .$$

A feladatot fogalmazzuk meg a

$$/2.16/ \quad \min f(x)$$

$$/2.17/ \quad g_i(x) \geq 0 \quad i = 1, \dots, m$$

alakban. Célunk a $g_1(x)$ feltétel kiértékelésével kapcsolatos problémák vizsgálata.

A $g_1(x)$ függvény több fontos gyakorlati eloszlás esetén logaritmikusan konkáv Prékopa András tétele alapján. Így ha az $f, -g_1, \dots, -g_m$ függvények konvexek, akkor a lo-

garitmikus büntetőfüggvény alkalmazása konvex segédfeladatra vezet. Ez volt több gyakorlati feladat megoldásának az útja (ld. [5]).

Sajnos a $g_1(x)$ függvény kiértékelése Monte-Carlo módszerrel igen munkaigényes. Sztohasztikus eljárás szerkesztéséhez pedig az $\ln g_1(x)$ függvényre kellene torzítatlan becslést adnunk, ami általánosságban nem látszik lehetségesnek.

Ha előre tudjuk, hogy $g_1(x) = 0$, akkor alkalmazhatunk egy $r^{-1}g_1^2(x)$ büntetőtagot. A $g_1^2(x)$ függvényre már kaphatunk torzítatlan becslést:

$$/2.18/ \quad g_1^2(x) = M(\xi_1(x) \cdot \xi_2(x)),$$

ahol $\xi_1(x), \xi_2(x)$ a $\xi(x)$ valószínűségi változó két független realizációját jelöli. Ugyanakkor azonban nincs garancia arra, hogy a logaritmikus és négyzetes tagokat is tartalmazó vegyes büntetőfüggvény konvex. Ugy tűnik tehát, hogy egy jó kezdeti közelítés kiszámítása nehéz feladat.

Tegyük fel azonban, hogy az aktiv indexeket már sikerült kijelölni és adaptáljuk a multiplikátormódszert. Legyen a feladat

$$/2.19/ \quad \min f(x)$$

$$/2.20/ \quad h_j(x) = 0 \quad j = 1, \dots, p.$$

Itt is mint előbb $h_j(x) = g_j(x)$ $j=1, \dots, p$ -re.

A bővített Lagrange-függvényt jelölje $Q(x, w, k)$ és vezessük be a

$$\begin{aligned} /2.21/ \quad \hat{Q}(x, w, k) &= f(x) + \sum_{j=2}^p w_j h_j(x) + k \sum_{j=2}^p h_j^2(x) + \\ &w_1 \xi_1(x) + k \xi_1(x) \xi_2(x) \end{aligned}$$

véletlen függvényt. Világos, hogy

$$/2.22/ \quad M(\hat{Q}(x, w, k)) = Q(x, w, k).$$

Igy $Q(x, w, k)$ x -szerinti $x(w)$ minimumának a meghatározása történhet a Kiefer-Wolfwitz módszerrel.

Tegyük fel, hogy a $g_1(x)$ függvény gradiensére is van torzítatlan becslésünk:

$$/2.23/ \quad \nabla g_1(x) = \nabla h_1(x) = M(\zeta(x)).$$

Ez a helyzet, ha $g_1(x)$ a /2.14/ formulával van definiálva és a parciális deriváltakat adó integrálkifejezéseket Monte-Carlo módszerrel értékeljük ki.

Ekkor a

$$\begin{aligned} /2.24/ \quad \nabla_x \hat{Q} &= \nabla_x f(x) + \sum_{j=2}^p w_j \nabla_x h_j + 2k \sum_{j=2}^p h_j(x) \nabla_x h_j(x) + \\ &+ w_1 \zeta(x) + 2k \xi_1(x) \zeta(x) \end{aligned}$$

véletlen vektor a $\nabla_x Q$ vektor torzítatlan becslése, feltéve, hogy a $\xi_1(x)$, $\zeta(x)$ valószínűségi változók függetlenek:

$$/2.25/ \quad \nabla_x Q = M(\nabla_x \hat{Q}).$$

Ekkor pedig a Robbins-Monroe típusu eljárás alkalmazható a $\nabla_x Q(x, w, k) = 0$ egyenletrendszer megoldására.

Most megmutatjuk, hogy a III. fejezetbeli /4.13/ eljárás sztohasztikus változata is könnyen kapható.

Vezessük be a

$$/2.26/ \quad \hat{H}(x) = (\zeta(x), \nabla g_2(x), \dots, \nabla g_m(x))$$

mátrixot. Világos, hogy a $\hat{H}(x)$ mátrix a /IV.2.11/-beli $H(x)$ mátrix torzítatlan becslése:

$$/2.27/ \quad H(x) = M(\hat{H}(x)).$$

Vezessük be a

$$/2.28/ \quad \hat{h}(x) = (\xi(x), h_2(x), \dots, h_p(x))'$$

véletlen vektort. Ez a $h(x) = (h_1(x), \dots, h_p(x))'$ vektor torzítatlan becslése. Végül legyen

$$/2.29/ \quad \hat{\delta w} = 2k\hat{h}(x)$$

$$/2.30/ \quad \hat{\delta x} = -2k\hat{G}\hat{H}(x)\hat{h}(x) - \hat{G}\nabla Q$$

Világos, hogy

$$/2.31/ \quad \delta w = M(\hat{\delta w})$$

és

$$/2.32/ \quad \delta x = M(\hat{\delta x})$$

ezért a 35. tétel alapján egy sztohasztikus eljárás szerkeszthető.

Newton-típusú eljárások esetén lényeges nehézséget okoz a $(\nabla_{xx}^2 Q)^{-1}$ mátrixra torzítatlan becslést adni. Egy dimenzió esetén arról van szó, hogy olyan ψ valószínűségi változót keresünk, amelyre

$$\frac{1}{m} = M(\psi)$$

ahol

$$m = M(\varphi) .$$

Itt φ ismert /realizálható/ valószínűségi változó és ψ -t φ realizációi alapján kell konstruálni. Ez tudomásom szerint általánosságban nem megoldható feladat. Mégis lehetséges Newton-típusú módszer sztohasztikus változatának a kidolgozása, bár ezek gazdaságossága nem nyilvánvaló. Erre a kérdésre még visszatérünk a Függelékben.

V. FÜGGELÉK

1. Egy dekompozíciós eljárás.

A dolgozatban több helyen felmerült összetett eljárások egyszerűsítésének a kérdése. A szóbanforgó esetekben az algoritmus konvergenciáját Ljapunov-függvényekkel állapíthatjuk meg. Ezért ezzel az esettel foglalkozunk általában is. Az algoritmusok folytonos alakban egy differenciálegyenlettel írhatók le:

$$\dot{x}(t) = \tilde{p}_1(x(t)).$$

Az algoritmus konvergenciája azt jelenti, hogy egy keresett x^* megoldás az egyenlet stabil egyensúlyi pontja, azaz $x(t) \rightarrow x^*$ $t \rightarrow \infty$ esetén, ha csak $x(0)$ elegendő közel van x^* -hoz.

Az egyenlet jobboldalán álló kifejezésben egyes részeket $Y(x)$ -szel jelölünk, és az egyenletet az

$$/1.1/ \quad \dot{x} = p_1(x, Y(x))$$

alakban írjuk át. Az $Y(x)$ kifejezés kijelölése lehet magától értetődő, de lehet mesterkelt is. Például a Newton-módszer alkalmazásakor az $\dot{x} = -f_{xx}^{-1} f_x$ egyenletben kiválaszthatjuk az $Y(x) = f_{xx}^{-1} f_x$ kifejezést.

Vagy például egy feltételes minimalizálási feladat lehet a

$$\begin{aligned} \min f(x, y) \\ g_i(x, y) = 0 \quad i = 1, \dots, m \end{aligned}$$

alakban megfogalmazva, ahol y komponenseinek a száma m . Ha y -t mint x függvényét kifejezzük a feltételrendszerből, akkor a

$$\min f(x, Y(x))$$

feladatot kell megoldani. /Ez az eljárás a redukált gradiens módszer./

Az $Y(x)$ kifejezés kiválasztását az indokolja, hogy ennek kiértékelése csak algoritmikus úton lehetséges. A megfelelő algoritmust jelölje

$$/1.2/ \quad \dot{y}(t) = r(x, y(t))$$

Az /1.2/ differenciálegyenlet jobboldalán x paraméterként szerepel, és $y(t) \rightarrow Y(x)$ ha $t \rightarrow \infty$ és $y(0)$ elég közel van $Y(x)$ -hez. A kérdés az, hogyan lehetne az /1.1/ ill. /1.2/ egyenletekkel definiált algoritmusokat egyesíteni, hogy elkerüljük $Y(x)$ pontos kiértékelését.

Egy lehetséges eljárás a következő: tetszőleges x, y pont esetén, amely az $x^*, Y(x^*)$ pont elég jó közelítése, definiáljuk a következő irányokat:

$$/1.3/ \quad p_1 = p(x, y), \quad r = r(x, y), \quad p_2 = A(x, y)p_1(x, y)$$

ahol

$$/1.4/ \quad A = A(x, y) = - r_x r_y^{-1}$$

Az A deriváltmátrixot az

$$/1.5/ \quad r(x, Y(x)) = 0$$

feltétel x szrinti deriválásával kapjuk.

Az algoritmust az

$$/1.6/ \quad \dot{x} = p_1(x, y)$$

$$/1.7/ \quad \dot{y} = p_2(x, y) + r(x, y)$$

differenciálegyenlettel definiáljuk. A következőkben az /1.6/ /1.7/ algoritmus konvergenciájával foglalkozunk.

Vezessük be a

$$z = (x, u)', \quad p = (p_1', p_2')'$$

jelöléseket. A Ljapunov-féle stabilitáselmélet szerint az /1.2/ differenciálegyenlet az $Y(x)$ pontban stabil, akkor és csak akkor, ha létezik egy folytonosan differenciálható $V(x, y)$ Ljapunov-függvény, amelyre

$$V(x, Y(x)) = 0, \quad V(x, y) > 0 \quad \text{ha } y \neq Y(x)$$

és

$$r' \nabla_y V < 0 \quad \text{ha } r \neq 0.$$

A $V(x, y)$ függvény az $Y(x)$ pont egy környezetében van értelmezve. Megköveteljük, hogy az erősebb

$$/1.8/ \quad r' \nabla_y V < - \varepsilon \|r\| \|\nabla_y V\|$$

egyenlőtlenség teljesüljön és azt mondjuk, hogy az /1.2/ differenciálegyenlet szigorúan stabil $Y(x)$ -ben. Ekkor V választható a

$$/1.9/ \quad V = r' K r$$

alakban, ahol K egy pozitív definit szimmetrikus mátrix. Az /1.8/ feltétel ekkor az

$$/1.10/ \quad r' (r_y K + K r_y') r < - \varepsilon \|r\|^2$$

alakban írható fel. Mivel az r irány az m -dimenziós y -nal koordinátázott altér tetszőleges iránya lehet, azért következik, hogy az $(x, Y(x))$ pontban az

$$/1.11/ \quad R(r_y K) = r_y K + K r_y'$$

mátrix pozitív definit. Ez a megfogalmazás is Ljapunovtól származik.

Folytonossági megfontolásból nyilvánvaló, hogy ha x az x^* egy kis környezetében fut végig, akkor a K mátrix x -től függetlennek is választható.

Tekintsük most a

$$/1.12/ \quad \dot{z} = p(z)$$

egyenletet. A kiindulást képező /1.1/ egyenlettel való kapcsolata a következő:

Rögzítsünk egy c konstans vektort és tekintsük az

$$/1.13/ \quad r(x, Y(x, c)) = c$$

egyenlettel definiált $Y(x, c)$ un. Ljapunov-felületet.

Az /1.1/ differenciálegyenlettel párhuzamosan tekintsük az

$$/1.14/ \quad \dot{x} = p_1(x, Y(x, c))$$

differenciálegyenleteket.

Tegyük fel, hogy az /1.1/ differenciálegyenlet szigorúan stabil x^* -ban, ekkor az /1.14/ differenciálegyenlet szintén szigorúan stabil valamilyen $x^*(c)$ pontban, amelyet a

$$/1.15/ \quad p_1(x^*(c), Y(x^*(c), c)) = 0$$

egyenlet definiál, ha csak c elég közel van az 0-hoz. Az $(x^*(c), Y(x^*(c), c))$ pontok egy m -dimenziós felületet töltenek ki, amelynek az $Y(x)$ felülettel egyetlen közös pontja van; $(x^*, Y(x^*))$.

Az /1.13/ egyenletből világos, hogy a $p = (p'_1, p'_2)'$ vektor bármely pontban egy azon a ponton áthaladó Ljapunov-felületnek érintője, ezért az /1.14/ differenciálegyenlet bármely megoldása teljes egészében egy $(x, Y(x, c))$ felületen

halad végig. Ezt a felületet tetszőleges módon /pl. az x -szel/ paraméterezve alkalmazhatjuk Ljapunov eredményeit és a következőt kapjuk:

Létezik egy H szimmetrikus mátrix, amely pozitív definit az $(x, Y(x(c)))$ felület $(x^*(c), Y(x^*(c), c))$ -ban vett $T(c)$ érintőterén, és az

$$/1.16/ \quad U = p'Hp$$

képlettel kiszámított függvény Ljapunov-függvénye az /1.14/ egyenletnek a következő értelemben:

$$p' \nabla_z U < - \varepsilon \|p\| \|\nabla_z U\|.$$

Ez úgy is fogalmazható, hogy

$$/1.17/ \quad p'(p_z H + H p'_z) p < - \varepsilon \|p\|^2$$

$$\text{ha} \quad p \in T(c).$$

Világos továbbá, hogy a p irányban a V függvény iránymenti deriváltja 0, mivel V konstans az $(x, Y(x, y))$ felületen:

$$/1.18/ \quad p' \nabla_z V = 0.$$

Ezen előkészítés után könnyű bebizonyítani a következőt
36. Tétel. Tegyük fel, hogy teljesülnek az /1.9 /- /1.17/ feltételek. Ekkor $z^* = (x^*, Y(x^*))$ az /1.6/ /1.7/ differenciálegyenlet szigorúan stabil egyensúlyi pontja, és

$$/1.19/ \quad W = U + \lambda V$$

az /1.6/ /1.7/ egyenlet Ljapunov-függvénye, ha csak λ elég nagy pozitív szám.

Bizonyítás: Vezessük be az $\tilde{r} = (0, r)$ jelölést, ahol 0 egy n dimenziós x -szel azonos dimenziójú 0 vektort jelent. Ekkor az /1.6/ /1.7/ differenciálegyenlet a

$$/1.20/ \quad \dot{z} = p(z) + \tilde{r}(z) = s(z)$$

alakban írható. A z^* pont az /1.20/ egyenlet egyensúlyi pontja és $W(z^*) = 0$, továbbá $W(z) > 0$ ha $z \neq z^*$.

A Finsler-lemma segítségével könnyű megmutatni, hogy W a z^* pont környezetében konvex. Megmutatjuk, hogy W az $s(z)$ irányban csökkenő. Valóban

$$\begin{aligned} /1.21/ \quad s(z)' \nabla_z W &= (p(z) + \tilde{r}(z))' (R(p_z H) p + \lambda R(\tilde{r}_z K) \tilde{r}) = \\ &= p' R(p_z H) p + \tilde{r}(z)' R(p_z H) p + \lambda \tilde{r}' R(r_z K) \tilde{r}. \end{aligned}$$

Figyelembe vettük, hogy /1.18/ alapján

$$p' R(\tilde{r}_z K) \tilde{r} = 0.$$

Az /1.21/ egyenlőség jobboldalára alsó becslést ad egy

$$/1.22/ \quad -a \|p\|^2 - b \|r\| \|p\| - c \lambda \|r\|^2$$

kvadratikus alak, ahol a és c pozitívak. Világos, hogy ha λ elég nagy, akkor az /1.22/ kvadratikus alak pozitív definit és ezzel a tételt bebizonyítottuk.

A 36. tétel alkalmazásaként megmutatjuk, hogyan lehet a redukált gradiens módszert egyszerűsíteni.

Tekintsük a

$$/1.23/ \quad \min f(x, y)$$

$$/1.24/ \quad g_i(x, y) = 0 \quad i = 1, \dots, m$$

feladatot, ahol y komponenseinek a száma m . Az /1.24/ feltételek által definiált $Y(x)$ függvényt az

$$/1.25/ \quad y = - Gg(x, y)$$

/folytonos alakban felírt/ algoritmus alapján számítjuk ki, ahol $g = (g_1, \dots, g_m)'$ és G az /1.24/ egyenlet Jacobi-mátrixa inverzének egy közelítése.

A

$$/1.26/ \quad \varphi(x) = f(x, Y(x))$$

függvényt gradiens módszerrel minimalizálva az

$$/1.27/ \quad \dot{x} = - f_x - Af_y$$

algoritmust kapjuk, ahol

$$A = \frac{\delta Y}{\delta x} = - g_x g_y^{-1}$$

Az /1.25/ ill. /1.27/ algoritmusok kompozíciójából az

$$/1.28/ \quad \dot{x} = - f_x - Af_y$$

$$/1.29/ \quad \dot{y} = - Gg - A'(f_x + Af_y)$$

algoritmust kapjuk.

Látható módon nem tudjuk elkerülni, hogy a g_y^{-1} mátrixot pontosan kiértékeljük, ez az A mátrix kiszámításához kell. Ez a redukált gradiens módszernek kétségtelenül hátránya a multiplikátormódszerrel szemben.

2. Sztohasztikus Newton-módszer.

Tekintsük a

$$\min f(x)$$

feladatot és alkalmazzuk a Newton-módszert. Folytonos alakot használva az

$$/2.1/ \quad \dot{x} = -(\nabla_{xx}^2 f)^{-1} \nabla_x f$$

algoritmust kapjuk. A jobboldali kifejezésben bevezetjük az

$$/2.2/ \quad Y(x) = (\nabla_{xx}^2 f)^{-1}$$

jelölést és $Y(x)$ -et iteratív módon határozzuk meg. Egy nonszinguláris A mátrix $Y^* = A^{-1}$ inverzét az

$$/2.3/ \quad \dot{Y} = -Y(AY-E)$$

algoritmus alapján határozhatjuk meg. Megjegyezzük, hogy az

$$/2.4/ \quad Y_{n+1} = 2Y_n - Y_n A Y_n$$

diszkrét változat kvadratikusan konvergens.

A /2.3/ algoritmusban A, Y független változóknak tekinthetők, és azt vizsgáljuk, hogy az A perturbációja milyen perturbációt idéz elő az Y -ban. Az /1.13/ képlet alapján az $r(x, y) = c$ Ljapunov-felületnek most az $Y(AY-E) = c$ felületek felelnének meg. A δA korrekciónak megfelelő δY korrekció kiszámításakor azonban az A^{-1} mátrixot ismernünk kellene.

Ezért a Ljapunov-felületeket az

$$/2.5/ \quad Y^{-1} - A = C$$

egyenlettel definiáljuk, ahol C konstans mátrix.

Tegyük fel, hogy A a t idő folytonos függvénye. Ekkor t szerinti differenciálással kapjuk az

$$/2.6/ \quad - Y^{-1} \dot{Y} Y^{-1} - \dot{A} = 0$$

egyenletet, ahonnan

$$/2.7/ \quad \dot{Y} = - Y \dot{A} Y.$$

Az előzőekben ismertetett /1.6/ /1.7/ konstrukció most a következőt adja:

$$/2.8/ \quad \dot{x} = - Y \nabla_x f$$

$$/2.9/ \quad \dot{Y} = - Y(A(x) Y^{-1} E) - Y \frac{d}{dt} A(x(t)) Y$$

ahol

$$A(x) = \nabla_{xx}^2 f(x).$$

A /2.8/ /2.9/ eljárás a Newton-módszernek egy módosítása, amelyhez hasonlókat már korábban is szerkesztettek. Ilyen pl. a Moser-féle eljárás. Az új eljárásunk egy lényeges vonása, hogy annak sztohasztikus változata is kidolgozható. Valóban, legyen $f(x)$ egy x -től, mint paramétertől függő többdimenziós integrál. Ekkor az f deriváltjaira Monte-Carlo módszerrel torzítatlan becslés adható, mondjuk

$$\nabla_x f = M(\xi(x)),$$

$$\nabla_{xx}^2 f = A(x) = M(\hat{A}(x)),$$

és

$$\nabla_{xxx}^3 f = M(B(x)).$$

Itt $\xi(x)$, $\hat{A}(x)$, $B(x)$ x -től függő realizálható valószínűségi változók. A $\nabla_{xxx}^3 f$ kifejezés egy háromindexes tenzort jelent. A /2.8/, /2.9/ algoritmus sztohasztikus változatát úgy kapjuk, hogy a jobboldalakat azok valamely torzítatlan becslésével helyettesítjük. Így kapjuk az

$$/2.10/ \quad \dot{x} = -Y \xi(x)$$

$$/2.11/ \quad \dot{Y} = -Y(\hat{A}(x)Y - E) - Y(B(x) \times Y \xi(x))Y$$

eljárást, amely Ivankov tétele alapján konvergens.

A /2.11/ kifejezésben szereplő \times egy háromindexes és egy egyindexes tenzor szorzatát jelöli. A torzítatlansághoz fel kell tennünk, hogy B ill. ξ függetlenek.

A /2.10/, /2.11/ algoritmust sztohasztikus Newton-módszernek nevezzük. A szerkesztés célja az volt, hogy az aránytalanul időigényes Monte-Carlo módszerről a "terhelést" a külső külső ciklusnak /jelen esetben a Newton-módszernek/ adjuk át. Érdekes volna megvizsgálni, hogy a harmadrendű deriváltakat tartalmazó tagok elhagyhatók-e.

IRODALOMJEGYZÉK

- [1] Armacost, R.L. and Fiacco, A.V., Computational experience in sensitivity analysis for non-linear programming. Mathematical Programming 6/1974/ No. 2, 301-326.
- [2] Bartels, R.H. Golub, R.H. and Saunders, M.A., Numerical Techniques in Mathematical Programming. In "Nonlinear Programming" /Ed. Rosen, J.B. Mangasarian, O.L. and Ritter, K./ Academic Press, New York - London, 1970, pp. 123-176.
- [3] Balakrishnan, A.V., On a new computing technique in optimal control. SIAM J, Control, 6/1968/ No.2, 149-173.
- [4] Belenkij, V.Z., Volonszkij, V.A., Ivankov, Sz.A., Pomanszkij, A.B., Sapiro, A.D., Iterativnyje metodü v tyeorii igr i programmirovanyii. Nauka, Moszkva 1974.
- [5] Deák István, Egy sztohasztikus programozási modell számítógépes kiértékelése. MTA SZK. Közlemények 1972. 9. 33-49.
- [6] Deák István: A többdimenziós tér halmazi valószínűségeinek kiszámítása normális eloszlás esetén. Alk. Mat. Lapok. 3/1976 /sajtó alatt/.
- [7] Fenchel, W., Über konvexe Funktionen mit vorgeschriebenen Niveaumannigfaltigkeiten. Math.Z. 63/1956/, 496-506.
- [8] Ferland, I.A., Mathematical Programming problems with quasi-convex objective functions. Mathematical Programming, 3/1972/, 296-301.

- [9] Fiacco, A.V. and McCormick, G.P., Nonlinear programming: Sequential Unconstrained Minimization Techniques. Wiley, New-York-London. 1968.
- [10] Fletcher, R., An exact penalty function for nonlinear programming with inequalities. Mathematical Programming 5/1973/, 129-150.
- [11] Forgó Ferenc, Egy módszer nemlineáris programozási feladatok közelítő megoldására. Szigma, I. 1969. 1. 67-75 old.
- [12] Fox, L., An Introduction to Numerical Linear Algebra. Clarendon Press, Oxford, 1964.
- [13] Gerencsér László: A készletgazdálkodás matematikai módszerei. Tanulmány az ELTE részére. 1972.
- [14] Gerencsér László, Extension of the extrapolation technique in the SUMT method. Problems of Control and Information Theory, pp 311-316./1975/.
- [15] Gerencsér László, Csizslennüje metodü v tyeorii igr. ZSVM 15/1975/N^o 3, 608-614 old.
- [16] Gerencsér László, A decomposition method in nonlinear programming. Math. Operationsforschung 6/1975/ Heft 4, S. 549-559.
- [17] Gerencsér László, A Second Order Technique for the Solution of Nonlinear Optimization Problems. Proc. of the Int. Conf on.O.R. held in Eger, 1974. "Progress in Operations Research". Coll. Math.Soc. János Bolyai, 12.
- [18] Gerencsér László - Bernau Heinz-Mayer János-Rapcsák Tamás, A nemlineáris programozás statisztikai alkalmazásai. Tanulmány az ELTE részére, 1975.

- [19] Grossmann, C., Eine neue parameterfreie Strafmethode in der nichtlinearen Optimierung Wiss.Z.TH. Ilmenau, 18/1972/ 3, 159-178.
- [20] Kéri Gerzson, An examination of nonnegativity and quasiconvexity conditions of quadratic forms on the nonnegative orthant. Studia Sci. Mat. Hung. 7/1972/, 11-20.
- [21] Klafszky Emil. Geometriai programozás és néhány alkalmazása. /Kandidátusi értekezés/ SZTAKI Tanulmány sorozat 1973/8 /
- [22] Kovács László Béla: Gradient Projection Method for Quasi-Concave Programming. Coll. az Appl. of Math. /szerk. Prékopa András/ 1965, 213-225 old.
- [23] Kerekó Béla: Optimumszámítás. Nemlineáris programozás. Közgazdasági és Jogi Könyvkiadó, 1972.
- [24] Martos Béla, Subdefinite Matrices and Quadratic Forms, SIAM J. Appl. Math. 17/1969/ 1215-1223.
- [25] Mayer János: Some computational experience with the reduced gradient method. Proc. Int. Conf. on Op. Res. Eger, 1974.
- [26] Mojszejev, N.N., Elementü tyeorii optimalnüh szisztjem. Nauka, 1975.
- [27] Prékopa András: Lineáris programozás, Bolyai J. Mat. Társ., 1968.
- [28] Prékopa András: Sztohasztikus rendszerek optimalizálási problémáiról. Akadémiai doktori disszertáció, Budapest, 1970.

- [29] Prékopa András: Contributions to the theory of stochastic programming, Mathematical Programming, /1973/, 202-221.
- [30] Prékopa András: Eine Erweiterung der sogenanter Methode der "zulässige Richtungen" der nicht-linearen Optimierung auf der Fall quasikonkaver Restriktionsfunktionen. 1973, Math. Operationsforschung u. Stat.
- [31] Rapcsák Tamás: Egy tározási modell számítástechnikai megoldása. Egyetemi doktori disszertáció 1974.
- [32] Rockafellar, R.T., A dual approach to solving nonlinear programming problems by unconstrained optimization. Mathematical Programming, 5/1973/No. 3. 354-374.
- [33] Strazicky Beáta, Some Results, concerning, an Algorithm for the Solution of the Two-Stage Stochastic Programming Problem. Proc.Int.Conf. on Stoch. Progr., /ed. M. Dempster/ Oxford, 1974.
- [34] Szidarovszky F., Bevezetés a numerikus módszerekbe. Közgazdasági és Jogi Könyvkiadó, 1974.
- [35] Wasan, M.T.: Stochastic approximation, Cambridge Univ.Press, 1969.

A TANULMÁNYOK sorozatban eddig megjelentek:

- 1/1973 Pásztor Katalin: Műszerek Boole-függvények minimális vagy nem redundáns, $\{\wedge, \vee, \neg\}$ vagy $\{\text{NAND}\}$ bázisneli, zárójeles vagy zárójel nélküli formuláinak előállítására
- 2/1973 Ванкеви Иштван: Расчленение многосвязных промышленных процессов с помощью вычислительных машин
- 3/1973 Ádám György: A számítógépipar helyzete 1972 második felében
- 4/1973 Bányász Csilla: Identification in the Presence of Drift
- 5/1973* Gyürki J. - Laufer J. - Grint M. - Somló J.: Optimalizáló adaptív szerszámgepirányítási rendszerek
- 6/1973 Szelke E. - Tóth K.: Felhasználói Kézikönyv /USER MANUAL/ a Folytonos Rendszerek Szimulációjára készült ANDISIM programnyelvhez
- 7/1973 Legendi Tamás: A CHANGE nyelv/multiprocesszor
- 8/1973 Klafszy Emil: Geometriai programozás és néhány alkalmazása
- 9/1973 R.Narasimhan: Picture Processing Pax
- 10/1973 Dibuz Á. - Gáspár J. - Várszegi S.: MANU-WRAP hátlaphuzalozó, MSI-TESTER integrált áramköröket mérő, TESTOMAT-C logikai hálózatokat vizsgáló berendezések ismertetése
- 11/1973 Matolcsi Tamás: Az optimum-számítás egy új módszeréről
- 12/1973 Makroprocesszorok, programozási nyelvek. Cikkgyűjtemény az NJSzT és SzTAKI közös kiadásában. Szerkesztette: Legendi Tamás
- 13/1973 Jedlovsky Pál: Új módszer bonyolult retifikáló oszlopok vegyész mérnöki számítására

- 14/1973 Bakó András MTA kutatóintézeteinek bérszámfejtése számítógéppel
- 15/1973 Ádám György: Kelet-nyugati kapcsolatok a számítógépiparban
- 16/1973 Fidrich I. - Uzsoy M.: LIDI-72 listakezelő rendszer a Digitális Osztályon, 1972. évi változat
- 17/1974 Gyürki József: Adaptív termelésprogramozó rendszer /APS/ termelőműhelyek irányítására
- 18/1974 Pikler Gyula: MINI-számítógépes interaktív alkatrész-programiró rendszer NC szerszámgépek automatikus programozásához
- 19/1974 Gertler, J.-Sedlak, J.: Software for process control
- 20/1974 Vámos, T.-Vassy, Z.: Industrial Pattern Recognition Experiment - A Syntax Aided Approach
- 21/1974 A KGST I.-15-1.: "Diszkrét rendszerek automatikus vezérlése" c. témában 1973. februárban rendezett szeminárium előadásai
- 22/1974 Arató, M.-Benczur, A.-Krámli, A.-Pergel, J.: Stochastic Processes, Part I.
- 23/1974 Benkó S. - Renner G.: Erősen telített mágneskörök számítógépes tervezései módszere
- 24/1974 Kovács György - Franta Lászlóné: Programcsomag elektronikus berendezések hátlaphuzalozásának tervezésére
- 25/1974 Járdán R. Kálmán: Háromfázisu tirisztoros inverterek állandósult tranziens jelenségei és belső impedanciája
- 26/1974 Gergely Lózsef: Numerikus módszerek sparse mátrixokra
- 27/1974 Somló János: Analitikus optimalizálás

- 28/1974 Vámos Tibor: Tárgyfelismerési kísérlet nyelvi módszerekkel
- 29/1974 Móricz Péter: Vegyészmérnöki számítási módszerek fázis-egyensúlyok és kémiai egyensúlyok vizsgálatára
- 30/1974 Vassy, Z.-Vámos, T.: The Budapest Robot - Pragmatic Intelligence
- 31/1975 Nagy István: Frekvenciaosztásos középfrekvenciás inverterek elmélete
- 32/1975 Singer D.-Borossay Gy.-Koltai T.: Gázhálózatok optimális irányítása különös tekintettel a Fővárosi Gázművek hálózataira
- 33/1975 Vámos, T.-Vassy, Z.: Limited and Pragmatic Robot Intelligence
- Mérő, L.-Vassy, Z.: A Simplified and Fastened Version of the Hueckel Operator for Finding Optimal Edges in Pictures
- Галло В.: Программа для распознавания геометрических образов, основанная на лингвистическом методе описания и анализа геометрических структур
- 34/1975 László Nemes: Pattern Identification Method for Industrial Robots by Extracting the Main Features of Objects
- 35/1975 Garádi-Krámli-Ratkó-Ruda: Statisztikai és számítástechnikai módszerek alkalmazása kórházi morbiditási vizsgálatokban
- 36/1975 Renner Gábor: Elektromágneses tér számítása nagyhőmérsékletű anyagban
- 37/1975 Edgardo Felipe: Specification problems of a process control display

- 38/1975 Hajnal Andrásné: Nemlineáris egyenletrendszerek megoldási módszerei
- 39/1975^x A. Abd El-Sattar: Control of induction motor by three phase thyristor connections in the secondary circuit
- 40/1975 Gerhardt Géza: QDP grafikus interaktív szubrutinok a CDC 3300-GD'71 grafikus konfigurációra
- 41/1975 Arató M.-Benczur, A.-Krámli, A.-Pergel, J.: Stochastic Processes, Part II.
- 42/1975 Arató Mátyás: Fejezetek a matematikai statisztikából számítógépes alkalmazásokkal
- 43/1975 Matavovszky Tibor - dr. Pásztorné Varga Katalin: Programrendszer Boole-függvényrendszer együttes egyszerűsítésére vagy minimalizálására
- 44/1975 Bacsó Nándorné: Pneumatikus áramköri hazardok
- 45/1975 Varga András: Ellenpárhuzamos félvezetőpárokkal vezérelt aszinkronmotoros hajtások számítási módszerei
- 46/1976 Galántai Aurél: Eglépéses módszerek lokális hibabecslései
- 47/1976 Abaffy József: A feltétel nélküli függvényminimalizálás kvadratikus befejezésű módszerei
- 48/1976 Strehó Mária: Stiff típusú közönséges differenciálegyenletek megoldásáról

A ^x -gal jelölt kivételével a sorozat kötetei megrendelhetők az Intézet könyvtáránál /Budapest, XIII. Victor Hugo u. 18-22./

