tanulmányok

47/1976

MTA Számitástechnikai és Automatizálási Kutató Intézet Budapest

MAGYAR TUDOMÁNYOS AKADÉMIA
SZÁMITÁSTECHNIKAI ÉS AUTOMATIZÁLÁSI KUTATÓ INTÉZETE

A FELTÉTEL NÉLKÜLI FÜGGVÉNYMINIMALIZÁLÁS KVADRATIKUS BEFEJEZÉSÜ MÓDSZEREI

Irta:

ABAFFY JÓZSEF

Tanulmányok 47/1976.

A kiadásért felelős:

Dr Vámos Tibor

ISBN 963 311 017 3

Készült az Országos Müszaki Könyvtár és Dokumentációs Központ
házi sokszorositójában
F.v.: Janoch Gyula

Tartalomjegyzék

BEVEZETÉS

A feltétel nélküli függvényminimalizálások az alkalmazott matematika számos területén, az irányitáselméletben, optimalizálási kérdéseknél, stb. fordulnak elő. A SUMT módszer révén az operációkutatással is szoros kapcsolatban áll [19].

A feltétel nélküli függvényminimalizálásokkal az ötvenes évek végén kezdtek el behatóan foglalkozni. 1959-ben jelent meg Davidon [13] cikke, amelynek továbbfejlesztése az eddig ismert leghatásosabb módszer, az ugynevezett Fletcher-Powell-Davidon [15] módszer (1963). 196o-ban Rosenbrock [4o] cikkével kezdődik meg a grádienst nem használó módszerek kidolgozása. 1961-ben Hooke és Jeeves [22] közölnek egy gradienst nem használó módszert. 1964-ben jelenik meg Powell [34] módszere, amelyet a gradienst nem használó módszerek között azóta is a legjobbnak tartanak. 1967-ben Broyden [7] már a Quasi-Newton módszerek egy osztályát határozta meg, és igazolta azok kvadratikus befejezését. Az első nem-szimmetrikus Quasi-Newton módszerek 1969-ben jelentek meg Pearson [32] cikkében. A jelenlegi kutatások főként a Quasi-Newton módszerekkel foglalkoznak. Ez indokolta, hogy kidolgozzunk egy olyan általános két paraméteres iterációs sémát, amelyből az ismert szimmetrikus és nem szimmetrikus Quasi-Newton módszerek a paraméterek alkalmas megválasztásával adódnak. Minthogy a megjelent módszerek száma igen nagy, ebben a dolgozatban csupán néhány un. kvadratikus befejezésü módszerrel foglalkozunk. Az ilyen módszerek használatát az indokolja, hogy ha egy minimalizálandó $f(\underline{x})$, $\underline{x} \in R^n$ valós értékü, konvex függvénynek ismerjük egy a minimumhelyhez elég közel eső $\underline{x}_o \in R^n$ kiinduló pontját, akkor az $f(\underline{x})$ függvény a minimum közelében kvadratikus függvénnyel jól közelithető.

Ebben a dolgozatban nem foglalkozunk azzal, hogy milyen feltevéseket kell tenni az $f(\underline{x})$ konvex függvényre ahhoz, hogy az egyes módszerek a minimum helyét meg tudják határozni.

A dolgozatban ismertnek tételezzük fel az egyváltozós vonalmenti minimalizáló módszereket. Az arany-metszés módszere [3o] , Fibonacci módszere [31] , stb. [6].

Az értekezés 3 fejezetből áll. Az I. fejezetben olyan két, gradiens mentes kvadratikus befejezésü módszert ismertetünk, amely a kvadratikus függvény konjugált irányainak meghatározásán alapul. Az I. fejezet önálló eredménye az, hogy bebizonyitjuk, hogy a Smith módszer [42] ekvivalens a Fox-Wilkinson módszerrel (1.9 tétel). Ezenkivül az I. fejezetben bevezetjük a költségfüggvény fogalmát. A költségfüggvény kvadratikus függvényekre vonatkozóan jól jellemzi az egyes módszerek jóságának mértékét.

A II. fejezetben a Quasi-Newton módszerekkel foglalkozunk. A II. fejezet önálló eredménye az, hogy megadunk egy általános iterációs sémát (3.23) - (3.27), amely két tetszőleges paramétert tartalmaz, és bebizonyitjuk ennek az 1.2 definició értelmében vett kvadratikus befejezését. (3.4 tétel) Az általános iterációs sémából levezethetők az ismert szimmetrikus és nemszimmetrikus Quasi-Newton módszerek, továbbá tartalmazza a Broyden által meghatározott szimmetrikus Quasi-Newton módszer osztályt és azonos a Huang féle 4 paraméteres osztállyal.

A III. fejezetben számitási eredményeket közlünk. Az ismertetett módszereket az MTA CDC 33oo-as számitógépre FORTRAN nyelven beprogramoztuk és azokat próbafüggvényeken lefuttattuk. A III. fejezet önálló eredménye a Quasi-Newton módszerek egy gyorsitási lehetősége, amely a gyakorlatban jól bevált. (2. pont).

1. fejezet

Gradiensmentes kvadratikusan konvergens módszerek.

Ebben a fejezetben olyan módszerekkel foglalkozunk, amelyek nem használják a minimalizálandó

(1.1) $f(\underline{x}) = \frac{1}{2}\underline{x}^T A\underline{x} + \underline{x}^T\underline{b} + c \qquad \underline{b},\underline{x} \in R^n,\ c \in R^1,$ A nxn-es

pozitiv definit szimmetrikus mátrix (a kvadratikus alak Hesse mátrixa), gradiensét. Az ilyen eljárásokra azért van szükség, mert általában a gradiens kiszámitása nehéz, illetve numerikusan instabil, lásd Tyihonov [45]. Ebben a fejezetben két módszert tárgyalunk, Smith [42] módszerét és Powell [34] algoritmusát. A Smith módszerről igazoljuk, hogy ekvivalens a lineáris algebrában használatos Fox-Wilkinson módszerrel.

1.1 Definició

A $\underline{q}_i \neq \underline{0}$, $\underline{q}_i \in R^n$, $i=1,2,\dots,n$ irányok A-konjugált irányok, ha

$$\underline{q}_i^T A\underline{q}_j = \begin{cases} \neq 0 & \text{ha } i \neq j \\ 0 & \text{ha } i = j \end{cases} \qquad (i,j = 1,2,\dots,n)$$

1.2 Definició

Azt mondjuk, hogy egy függvényminimalizáló séma kvadratikus befejezésü, ha egy n-változós pozitiv definit szimmetrikus mátrixszal rendelkező kvadratikus alak minimumát legfeljebb n-iterációs lépésben megadja.

Egy iterációs lépésen egy konjugált irány és a konjugált irány menti minimum meghatározását értjük.

A kvadratikus befejezésü szóhasználatot az irodalomban használatos, nem tul szerencsés, kvadratikusan konvergens kifejezés helyett használjuk.

Bevezetjük a költségfüggvény fogalmát.

1.3 Definició

A M függvényminimalizáló módszer költségfüggvénye az (1.1) kvadratikus alakra vonatkozóan:

$$1.2 \qquad K_M(\ell_i(\varepsilon),\ \omega_i(\varepsilon)) = \sum_{i=1}^{m} (\ell_i(\varepsilon) + n\ \omega_i(\varepsilon)) \qquad (m \le n)$$

ahol $\ell_i(\varepsilon)$ az iterációs-lépésenkénti függvényértékek kiszámitásának a száma, $\omega_i(\varepsilon)$ pedig az iterációs-lépésenkénti gradiens irány kiszámitásának száma. Nyilván mind ℓ_i és ω_i függ ε-tól, a lineáris minimalizálás kilépési feltételétől. Az m a szükséges iterációs lépések számát jelöli.

A költségfüggvénynek ez a definiciója természetes. Ugyanis általában nem a függvényminimalizáló módszer, hanem a benne előforduló függvényértékek kiszámitása igényel sok müveletet, különösen bonyolult függvények esetén.

A függvényminimalizálás irodalmában a gradiens vektor kiszámitását azonosnak tekintik n függvényérték kiszámitásával, mert a függvény gradiens egyetlen komponensének kiszámitása közelitőleg azonos müveleti igényü a függvény helyettesitési értékének kiszámitásával.

A költségfüggvény fenti definiciója ugyanakkor a vonalmenti minimalizálások számára vonatkozóan is ad információt. Példaként tekintsük a következőt.

Legyen a M módszer olyan, amely nem használja (1.1) gradiensét. ($\omega_i(\varepsilon) \equiv 0$). Legyen az iterációs-lépésenkénti vonalmenti minimalizálások száma V és egy vonalmenti minimalizálás átlagosan s függvényérték kiszámitását (adott $\varepsilon > 0$ pontosságon belül) igényelje. Ekkor $\ell_i(\varepsilon) = V.s$, és

$$K_M(\ell_i, \omega_i) = \sum_{i=1}^{n} V.s = n.V.s = V_M.s$$

ahol $V_M = nV$ a M módszerbeli vonalmenti minimalizálások száma, azaz $K_M(\ell_i, \omega_i)$ tartalmazza a vonalmenti minimalizálások számát.

1.4 Definició

Az $A \subset R^n$ altér ($\dim A = m$) és $\underline{a} \in R^n$ vektor által meghatározott lineáris sokaságon (hipersik) a

$$M^m = \{\underline{x} + \underline{a} \mid \underline{x} \in A\}$$

halmazt értjük.

A Smith módszer kvadratikus befejezésének bizonyitásához szükségünk van a következő lemmákra.

1.5 Lemma

Ha $\underline{q}_i \neq \underline{0}$, $i = 1,2,\dots,m$ $m \leq n$ (1.1)-re vonatkozó A-konjugált irányok M^m-ben, akkor az M^m, m dimenziós sokaságban (1.1) minimumát ugy is megkaphatjuk, hogy minden $\underline{q}_i \neq \underline{0}$, $i = 1,2,\dots,m$ irány mentén csak egyszer számitunk vonalmenti minimumot.

Bizonyitás.

Ha $\underline{x}_0 \in M^m$ tetszőleges, akkor az $\underline{x}_m \in M^m$ minimum helyét (1.1)-nek, a következőképpen fejezhetjük ki a $\underline{q}_1, \underline{q}_2, \dots, \underline{q}_m$

irányok segitségével:

$$(1.3) \qquad \underline{x}_m = \underline{x}_o + \sum_{i=1}^{m} \alpha_i \underline{q}_i$$

ahol az $\alpha_i \in R^1$ számokat ugy kell meghatároznunk, hogy $\underline{x}_m \in M^m$ minimumhelye legyen (1.1)-nek. Felhasználva (1.3)-at a következőt irhatjuk, kihasználva a $\underline{q}_i$, $i = 1,2,\dots,m$ vektorok A-konjugáltságát.

$$(1.4) \qquad f(\underline{x}_m) = f(\underline{x}_o + \sum_{i=1}^{m} \alpha_i \underline{q}_i) =$$

$$= \sum_{i=1}^{m} \left[{}^1/_2\, \alpha_i^2\, \underline{q}_i^T A\, \underline{q}_i + \alpha_i\, \underline{q}_i^T (A\, \underline{x}_o + \underline{b}) \right] + f(\underline{x}_o)$$

amiből következik, hogy az α_i, $i=1,2,\dots,m$ konstansok egymástól függetlenül meghatározhatók a $\underline{q}_i \neq \underline{O}$, $i=1,2,\dots,m$ vonalmenti minimalizálások segitségével.

<u>1.6 Lemma</u>

Ha a $\underline{q}_i \neq \underline{O}$, $i=1,2,\dots,k$ A-konjugált irányok két különböző lineáris sokaságban $(M^s,\ M^t,\ n > s,\ t \geq k > 0,\ M^s \cap M^t = \emptyset)$ és $\underline{x}_o \in M^s$ illetve $\underline{x}_1 \in M^t$, (1,1) minimumhelyei a megfelelő lineáris sokaságban, akkor

$$\underline{q}_i^T A(\underline{x}_1 - \underline{x}_o) = 0 \qquad (i=1,2,\dots,k)$$

azaz a $\underline{q}_i \neq \underline{O}$, $i=1,2,\dots,k$ és az $\underline{x}_1 - \underline{x}_o \neq \underline{O}$ irányok A-konjugált irányok R^n -ben.

Bizonyitás.

A lineáris sokaság 1.4 definiciójából következik, hogy

$M^s \subset R^n$, $M^t \subset R^n$ tehát $\underline{q}_i \neq \underline{0}$, i=1,2,...,m A-konjugált irányok R^n -ben. Mivel M^s, M^t két különböző lineáris sokaság, azaz $M^s \cap M^t = \emptyset$, következik $\underline{x}_1 - \underline{x}_o \neq \underline{0}$ és $\underline{x}_1 - \underline{x}_o \in R^n$
Minthogy $\underline{x}_o \in M^s$ minimumhely

$$\frac{d}{d\alpha} f(\underline{x}_o + \alpha\, \underline{q}_j) = 0 \text{ és } \alpha = 0 \text{ minden } j=1,2,\ldots,k \text{ -ra}$$

és hasonlóan $\underline{x}_1 \in M^t$ -re

$$\frac{d}{d\alpha} f(\underline{x}_1 + \alpha\, \underline{q}_j) = 0 \text{ és } \alpha = 0 \text{ minden } j=1,2,\ldots,k\text{-ra}$$

Felhasználva (1.1)-et a következő 2k egyenletet kapjuk

$$\alpha\, \underline{q}_j^T A\, \underline{q}_j + \underline{q}_j^T (A\, \underline{x}_o + \underline{b}) = 0 \qquad (j=1,2,\ldots,k)$$

$$\alpha\, \underline{q}_j^T A\, \underline{q}_j + \underline{q}_j^T (A\, \underline{x}_1 + \underline{b}) = 0 \qquad (j=1,2,..,k)$$

Minthogy $\alpha = 0$ kivonva a $\underline{q}_j$, j=1,2,...,k irányra vonatkozó azonos indexü egyenleteket egymásból a következő k egyenletet kapjuk:

$$\underline{q}_j^T A(\underline{x}_1 - \underline{x}_o) = 0 \qquad (j=1,2,\ldots,k)$$

amivel állitásunkat bebizonyitottuk.

Smith módszere

Legyen $\underline{v}_1, \underline{v}_2, \ldots, \underline{v}_n$ az R^n egy bázisa, $\underline{q}_1 = \underline{v}_1$ és $\underline{x}_o \in R^n$ tetszőleges.
Határozzuk meg az $\underline{x}_o$ pontból kiindulva a $\underline{q}_1$ irány mentén $f(\underline{x})$ minimumát $(\underline{x}_1^{m,1})$. Ez azt jelenti, hogy a $\underline{q}_1$ által definiált $S^1 \subset R^n$ egydimenziós altérben meghatároztuk (1.1) minimumát (1 vonalmenti minimalizálás).

Lépjünk az $\underline{x}_1^{m,1}$ pontból a $\underline{v}_2$ irányban egy tetszőleges nem zérus lépést. Az igy kapott $\underline{x}_1$ ($\underline{x}_1 \neq \underline{x}_1^{m,1}$) pontból a $\underline{q}_1$ irány mentén határozzuk meg $f(\underline{x})$ minimumát ($\underline{x}_2^{m,1}$).

Az $\underline{x}_1^{m,1} \neq \underline{x}_2^{m,1}$ és az 1.6 lemma miatt

$$\underline{q}_1^T A(\underline{x}_2^{m,1} - \underline{x}_1^{m,1}) = 0$$

azaz a $\quad \underline{q}_2 = \underline{x}_2^{m,1} - \underline{x}_1^{m,1} \neq \underline{0}$ irány A-konjugált $\underline{q}_1$ -re és $\underline{q}_1$, $\underline{q}_2$ lineárisan függetlenek.

A $\underline{q}_1$, $\underline{q}_2$ irányok definiálják az $S^2 \subset R^n$ kétdimenziós alteret. Ebben az altérben az 1.5 Lemma alapján megkapjuk $f(\underline{x})$ minimumát ugy, hogy a $\underline{q}_1$, $\underline{q}_2$ irány mentén meghatározzuk az $\underline{x}_2^{m,1}$, $\underline{x}_2^{m,2}$ minimumhelyeket (2 vonalmenti minimalizálás).

Az általános lépést a következőképpen irjuk le.

Tegyük fel, hogy már meghatároztuk a $\underline{q}_1, \underline{q}_2, \ldots, \underline{q}_{i-1}$ lineárisan független páronként konjugált irányokat, az általuk definiált $S^1 \subset S^2 \subset S^3 \subset \ldots \subset S^{i-1}$ altereket és az $\underline{x}_{i-1}^{m,i-1}$ pontot, amely $f(\underline{x})$ minimumhelye az $S^{i-1} \subset R^n$ altérben. A $\underline{q}_i$ A-konjugált irányt a következőképpen határozzuk meg.

Tegyünk az $\underline{x}_{i-1}^{m,i-1}$ pontból egy nem nulla lépést a $\underline{v}_i$ irányban. Az igy kapott $\underline{x}_i$ pontból a $\underline{q}_1$ irány mentén határozzuk meg $f(\underline{x})$ minimumát ($\underline{x}_i^{m,1}$), majd az igy kapott $\underline{x}_i^{m,1}$ pontból a $\underline{q}_2$ irány mentén határozzuk meg $f(\underline{x})$ minimumát ($\underline{x}_i^{m,2}$)stb.

Utolsó lépésként az $\underline{x}_i^{m,i-2}$ pontból a $\underline{q}_{i-1}$ irány mentén meghatározva a minimumot az $\underline{x}_i^{m,i-1}$ pontot kapjuk. Minthogy $\underline{x}_{i-1}^{m,i-1}$ és $\underline{x}_i^{m,i-1}$ különböző lineáris sokaságbeli minimumhelyek, $\underline{x}_{i-1}^{m,i-1} \neq \underline{x}_i^{m,i-1}$ és

$$\underline{q}_j^T A(\underline{x}_{i-1}^{m,i-1} - \underline{x}_i^{m,i-1}) = 0 \qquad (j=1,2,\ldots,i-1)$$

Ez azt jelenti, hogy a $\underline{q}_i = \underline{x}_{i-1}^{m,i-1} - \underline{x}_i^{m,i-1}$ irány a $\underline{q}_1, \underline{q}_2, \ldots, \underline{q}_{i-1}$ irányokra A-konjugált és $\underline{q}_1, \underline{q}_2, \ldots, \underline{q}_i$ lineárisan függetlenek. Ezek az irányok kifeszitik az $S^i \subset R^n$ i-dimenziós alteret, amelyben $f(\underline{x})$ minimumhelyét az 1.5 lemma alapján ugy kapjuk meg, hogy az $\underline{x}_i^{m,i-1}$ pontból kiindulva a $\underline{q}_i$ irány mentén meghatározzuk $f(\underline{x})$ minimumát (i vonalmenti minimalizálás).

Az eljárás n-edik lépése után kapott $\underline{x}_n^{m,n}$ pont $f(\underline{x})$ minimumhelye az R^n térben, minthogy a $\underline{q}_1, \underline{q}_2, \ldots, \underline{q}_n$ A-konjugált vektorok kifeszitik az n-dimenziós teret és az $\underline{x}_n^{m,n}$ pontot az 1.5 lemma segitségével határoztuk meg.

Bebizonyitottuk tehát a következő tételt.

<u>1.7 Tétel</u>

Legyen $\underline{v}_1, \underline{v}_2, \ldots, \underline{v}_n$ az R^n egy bázisa. Az $f(\underline{x})$ kvadratikus függvény minimumhelyét $(\underline{x}_n^{m,n})$ n lépésben meghatározhatjuk a következőképpen:

(1.5) 1/ $\underline{q}_1 = \underline{v}_1 \quad , \quad \underline{x}_o \in R^n$ tetsz.

(1.6) $$f(\underline{x}_1^{m,1}) = \min_{\alpha \in R^1} f(\underline{x}_o + \alpha\, \underline{q}_1)$$

(1.7) 2/ $\underline{x}_i = \underline{x}_{i-1}^{m,i-1} + \beta_i\, \underline{v}_i \quad , \quad \beta_i \neq 0$ tetsz.

(1.8) $$\underline{x}_i^{m,0} = \underline{x}_i$$

(1.9) $$f(\underline{x}_i^{m,j}) = \min_{\alpha \in R^1} f(\underline{x}_i^{m,j-1} + \alpha\, \underline{q}_j)(j=1,2,\ldots,i-1)$$

(1.10) $$\underline{q}_i = \underline{x}_i^{m,i-1} - \underline{x}_{i-1}^{m,i-1}$$

(1.11) $$f(\underline{x}_i^{m,i}) = \min_{\alpha \in R^1} f(\underline{x}_i^{m,i-1} + \alpha\, \underline{q}_i)$$

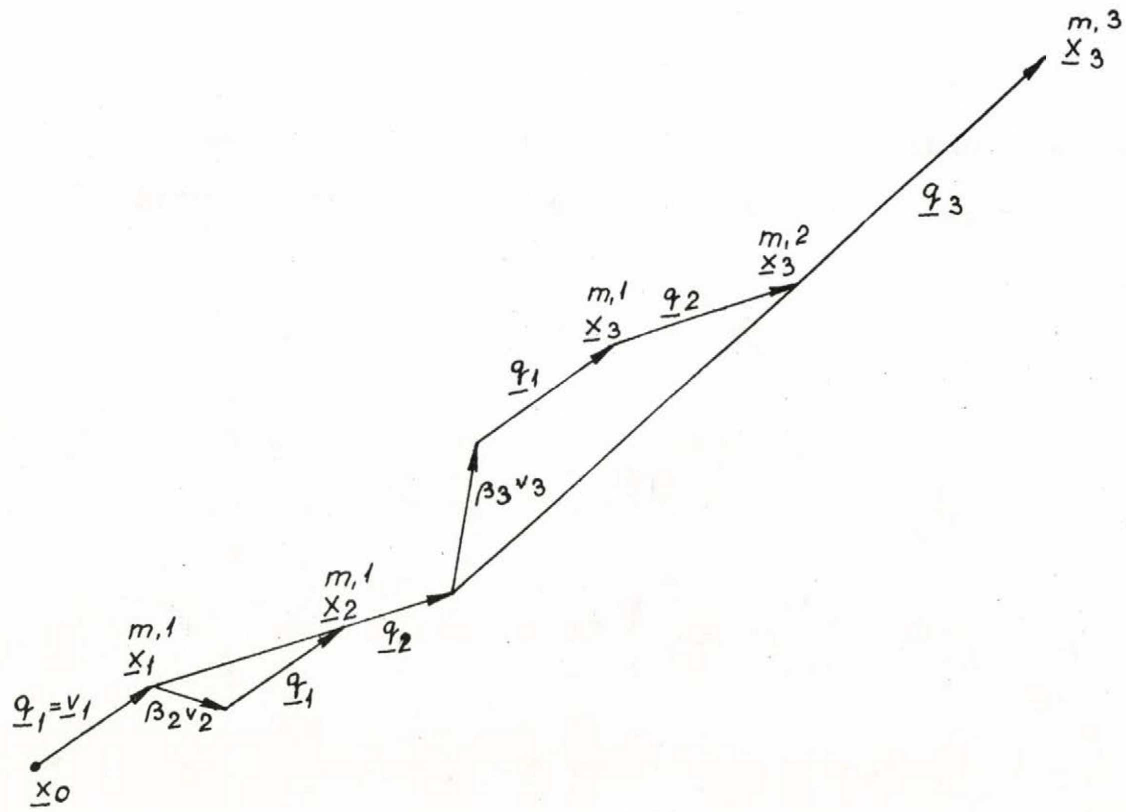
(i=2,3,n)

Az eljárás költségfüggvénye a tétel alapján

$$K_S = \frac{n(n+1)}{2}\, t$$

ahol t a vonalmenti minimalizálásnál kiszámitott függvényértékek száma.

Háromdimenziós esetben az eljárást az alábbi ábrával szemléltethetjük:



1. ábra

Kimutatjuk, hogy az (1.1) kvadratikus alak minimumának a Smith módszer szerinti meghatározása a Fox-Wilkinson féle [18] lineáris egyenletrendszerekre vonatkozó konjugált irányok módszerével ekvivalens. A Fox-Wilkinson módszerben egy tetszőleges bázisból kiindulva a konjugált irányok meghatározása és a lineáris egyenletrendszer megoldása $\frac{n(n+1)}{2}$ érték kiszámitását igényli. A Fox-Wilkinson szimmetrikus lineáris egyenletrendszert megoldó módszer a következő.

Legyen a megoldandó lineáris egyenletrendszer

(1.12) $A\,\underline{x} = -\,\underline{b}\,,\quad \underline{x} \in R^n\,,\quad A\;\; n \times n$ -es

pozitiv definit szimmetrikus mátrix és legyen

$$\underline{e}_1,\ \underline{e}_2,\ \dots,\ \underline{e}_n$$

az egységkoordináta irányokból képezett bázis R^n -ben.

Az A mátrixnak megfelelő konjugált irányokat a következőképpen kapjuk:

$$\underline{s}_1 = \underline{e}_1$$

(1.13)
$$\underline{s}_i = \underline{e}_i - \sum_{j=1}^{i-1} \gamma_{i,j}\,\underline{s}_j \qquad (i=2,3,\dots,n)$$

ahol

(1.14)
$$\gamma_{i,j} = \frac{\underline{e}_i^T A\,\underline{s}_j}{\underline{e}_j^T A\,\underline{s}_j} \qquad \begin{array}{l}(i=2,3,\dots,n)\\(j=1,2,\dots,i-1)\end{array}$$

A konjugált irányok segitségével az (1.12) $\underline{x}_m$ megoldása a következőképpen fejezhető ki

(1.15)
$$\underline{x}_m = \sum_{i=1}^{n} a_i\,\underline{s}_i = -\sum_{i=1}^{n} \frac{\underline{b}^T \underline{s}_i}{\underline{s}_i^T A\,\underline{s}_i}\,\underline{s}_i$$

ugyanis

(1.16)
$$a_i = -\frac{\underline{b}^T \underline{s}_i}{\underline{s}_i^T A\,\underline{s}_i} \qquad (i=1,2,\dots,n)$$

Először is megállapitjuk, hogy a Fox-Wilkinson módszer valóban $\frac{n(n+1)}{2}$ érték kiszámitását igényli, ugyanis a $\gamma_{i,j}$ együtthatók száma (1.14)-ből és az a_i száma (1.16)-ból

$$\frac{n(n-1)}{2} + n = \frac{n(n+1)}{2}$$

Legyen a Smith módszer induló bázisa szintén

$$\underline{e}_1, \underline{e}_2, \ldots, \underline{e}_n .$$

Ennek megfelelően a (1.5) - (1.11) kifejezésekben

$$\underline{v}_i \equiv \underline{e}_i \qquad (i=1,2,\ldots,n) \tag{1.17}$$

Legyen az $f(\underline{x})$ kvadratikus alak minimumhelye $\underline{x}'_m$.

Mivel (1.1)-ben A pozitiv definit mátrix az $f(\underline{x})$ függvény gradiens iránya akkor és csak akkor $\underline{0}$, ha $\underline{x} = \underline{x}'_m$

azaz

$$\text{grad } f(\underline{x}) = A\,\underline{x} + \underline{b} = \underline{0} \iff \text{ha} \quad \underline{x} = \underline{x}'_m$$

Innen $A\,\underline{x}'_m = -\underline{b}$, azaz $\underline{x}'_m$ megoldása (1.12)-nek, és forditva ha $\underline{x}_m$ megoldása (1.12)-nek akkor

$$\text{grad } f(\underline{x}_m) = \underline{0} .$$

Tehát $\underline{x}'_m \equiv \underline{x}_m$.

Megmutatjuk, hogy ha (1.17) szerint választjuk a Smith módszer bázisát, $\underline{x}'_o = \underline{0}$ az induló pont, akkor

$$\underline{q}_i \equiv \underline{s}_i \qquad (i=1,2,\ldots,n)$$

Legyen

$$\underline{s}_1 = \underline{q}_1 = \underline{e}_1 \tag{1.18}$$

Az (1.6) szerint az $\underline{x}'_o = \underline{0}$ pontból a $\underline{q}_1$ irány mentén kell a minimumot meghatározni

$$f(\underline{x}_1^{m,1}) = \min_{\alpha \in R^1} f(\underline{x}_o' + \alpha\, \underline{q}_1) = \min_{\alpha \in R^1} f(\alpha\, q_1).$$

Kihasználva az A mátrix pozitiv definitségét és szimmetrikus voltát

$$\frac{d\, f(\underline{x} + \alpha\, \underline{q}_i)}{d\, \alpha} = \underline{x}^T A\, \underline{q}_i + \alpha\, \underline{q}_i^T A\, \underline{q}_i + \underline{b}^T \underline{q}_i = 0 \tag{1.19}$$

$(i=1,2,\ldots,n)$

A $\underline{q}_i \neq \underline{0}$ $i=1,2,\ldots,n$ irány mentén (1.19) pontosan akkor teljesül, ha

$$\alpha = \frac{-\underline{b}^T \underline{q}_i - \underline{x}^T A\, \underline{q}_i}{\underline{q}_i^T A\, \underline{q}_i} \qquad (i=1,2,\ldots,n) \tag{1.2o}$$

Az (1.2o) relációt felhasználva

$$\alpha = -\frac{\underline{b}^T \underline{q}_1}{\underline{q}_1^T A\, \underline{q}_1} = -\frac{\underline{b}^T \underline{s}_1}{\underline{s}_1^T A\, \underline{s}_1} = a_1\,,$$

azaz az első vonalmenti minimalizálás, ami jelen esetben csak egy konstans kiszámitását jelenti (mivel az A mátrix és $\underline{b}$ ismeretes), adja az (1.15) kifejezés első tagjának konstansát. Az (1.6) szerint tehát az

$$\underline{x}_1^{m,1} = -\frac{\underline{b}^T \underline{q}_1}{\underline{q}_1^T A\, \underline{q}_1}\, \underline{q}_1 \tag{1.21}$$

pontból tetszőleges nem nulla eltolást kell végrehajtanunk az $\underline{e}_2$ irányban. Legyen most $\beta_2 = 1$ igy

$$\underline{x}_2^{m,o} = \underline{x}_2' = \underline{x}_1^{m,1} + \underline{e}_2$$

Ebből a pontból elvégezve az (1.9) szerinti minimalizálást, az (1.2o), (1.18) miatt

$$\alpha = \frac{-\underline{b}^T \underline{q}_1 - \underline{x}_2^{m,o^T} A \underline{q}_1}{\underline{q}_1^T A \underline{q}_1} = - \frac{\underline{e}_2^T A \underline{q}_1}{\underline{q}_1^T A \underline{q}_1} = - \frac{\underline{e}_2^T A \underline{s}_1}{\underline{e}_1^T A \underline{s}_1} = - \gamma_{2,1}$$

Tehát az (1.9) minimalizálásból megkaptuk a $\gamma_{2,1}$ konstanst

és az

$$\underline{x}_2^{m,1} = \underline{x}_2^{m,o} - \frac{\underline{e}_2^T A \underline{q}_1}{\underline{q}_1^T A \underline{q}_1} \underline{q}_1 \qquad \text{pontot.}$$

Az eljárást tovább folytatva kapjuk, hogy (1.lo) szerint

$$(1.22) \qquad \underline{q}_2 = \underline{x}_2^{m,1} - \underline{x}_1^{m,1} = \underline{x}_2^{m,o} - \frac{\underline{e}_2^T A \underline{q}_1}{\underline{q}_1^T A \underline{q}_1} \underline{q}_1 - \underline{x}_1^{m,1} =$$

$$= \underline{x}_1^{m,1} + \underline{e}_2 - \frac{\underline{e}_2^T A \underline{q}_1}{\underline{q}_1^T A \underline{q}_1} \underline{q}_1 - \underline{x}_1^{m,1} = \underline{s}_2$$

Tehát a $\underline{q}_2$ irány valóban egybeesik $\underline{s}_2$-vel. A minimumhely a $\underline{q}_2$ irány mentén az $\underline{x}_2^{m,1}$ pontból (1.2o) szerint, felhasználva (1.22)-t, a következő

$$\underline{x}_2^{m,2} = \underline{x}_2^{m,1} - \frac{\underline{b}^T \underline{q}_2 + \underline{x}_2^{m,1^T} A \underline{q}_2}{\underline{q}_2^T A \underline{q}_2} \underline{q}_2 = \underline{q}_2 + \underline{x}_1^{m,1} -$$

$$- \frac{\underline{b}^T \underline{q}_2}{\underline{q}_2^T A \underline{q}_2} \underline{q}_2 - \frac{\underline{q}_2^T A \underline{q}_2}{\underline{q}_2^T A \underline{q}_2} \underline{q}_2 - \frac{\underline{x}_1^{m,1^T} A \underline{q}_2}{\underline{q}_2^T A \underline{q}_2} \underline{q}_2 =$$

$$= - a_1 \underline{q}_1 - \frac{\underline{b}^T \underline{q}_2}{\underline{q}_2^T A \underline{q}_2} \underline{q}_2 = a_1 \underline{s}_1 + a_2 \underline{s}_2$$

ahol kihasználtuk, hogy $\underline{q}_1$, $\underline{q}_2$ A-konjugált irányok. Tehát az $\underline{x}_2^{m,2}$ pont az (1.15) kifejezés első két tagját adja, azaz ha a feladat egy 2x2 -es szimmetrikus, pozitiv definit mátrixszal rendelkező lineáris egyenletrendszer megoldása, akkor $\underline{x}_2^{m,2}$ a lineáris egyenletrendszer megoldását adja, és egyben a megfelelő $f(\underline{x})$ kvadratikus alak minimumhelyét. Vegyük észre, hogy az 1.7 kifejezésben a $\beta_i = 1$ $(i=1,2,\dots,n)$ választást kell tenni, továbbá az (1.9) kifejezés adja az $\gamma_{i,j}$ együtthatókat, az (1.11) minimalizálási feladat pedig az a_i együtthatókat.

Tegyük fel, hogy a k-adik lépésig eljutottunk, azaz

(1.23) $$\underline{q}_j = \underline{s}_j \qquad (j=1,2,\dots,k)$$

(1.24) $$\gamma_{i,j} = \frac{\underline{e}_i^T A \underline{s}_j}{\underline{e}_j^T A \underline{s}_j} \qquad (j=1,2,\dots,i-1),(i=1,2,\dots,k)$$

(1.25) $$\underline{x}_k^{m,k} = \sum_{i=1}^{k} a_i \underline{s}_i = \sum_{i=1}^{k} a_i \underline{q}_i , \quad a_i = - \frac{\underline{b}^T \underline{s}_i}{\underline{s}_i^T A \underline{s}_i}$$

$$(i=1,2,\dots,k)$$

és lássuk be, hogy

(1.26) $$\underline{q}_{k+1} = \underline{s}_{k+1}$$

(1.27) $$\gamma_{k+1,j} = \frac{\underline{e}_{k+1}^T A \underline{s}_j}{\underline{e}_j^T A \underline{s}_j} \qquad (j=1,2,\dots,k) ,$$

valamint

(1.28) $$\underline{x}_{k+1}^{m,k+1} = \sum_{i=1}^{k+1} a_i \underline{s}_i , \qquad a_{k+1} = - \frac{\underline{b}^T \underline{s}_{k+1}}{\underline{s}_{k+1}^T A \underline{s}_{k+1}} .$$

A $\beta_{k+1} = 1$ választással (1.7) -ből kapjuk, hogy

(1.29) $$\underline{x}_{k+1}^{m,o} = \underline{x}'_{k+1} = \underline{x}_k^{m,k} + \underline{e}_{k+1}$$

Végezzük el rendre az (1.9) szerinti minimalizálást (1.2o) szerint

$$(1.3o) \quad \underline{x}_{k+1}^{m,j} = \underline{x}_{k+1}^{m,j-1} - \frac{\underline{b}^T \underline{q}_j + \underline{x}_{k+1}^{m,j-1^T} A \underline{q}_j}{\underline{q}_j^T A \underline{q}_j} \underline{q}_j \quad (j=1,2,\dots,k)$$

Minthogy az (1.13) kifejezés az (1.23) indukciós feltevés miatt érvényes $\underline{q}_j$ -re $(j=1,2,\dots,k)$ igy az A-konjugáltságukat kihasználva kapjuk, hogy

$$(1.31) \quad \underline{q}_j^T A \underline{q}_j = \underline{e}_j^T A \underline{q}_j \quad (j=1,2,\dots,k)$$

Az (1.29), (1.31) kifejezés és az (1.3o) rekurziv képlet alapján $\underline{x}_{k+1}^{m,k}$ -ra a következőt kapjuk (lásd 1.8 lemma)

$$(1.32) \quad \underline{x}_{k+1}^{m,k} = \underline{x}_k^{m,k} + \underline{e}_{k+1} - \sum_{j=1}^{k} \frac{\underline{e}_{k+1}^T A \underline{q}_j}{\underline{e}_j^T A \underline{q}_j} \underline{q}_j$$

Az (1.1o) kifejezés és (1.32) alapján $\underline{q}_{k+1}$ -re a következő adódik

$$\underline{q}_{k+1} = \underline{x}_{k+1}^{m,k} - \underline{x}_k^{m,k} = \underline{e}_{k+1} - \sum_{j=1}^{k} \frac{\underline{e}_{k+1}^T A \underline{q}_j}{\underline{e}_j^T A \underline{q}_j} \underline{q}_j =$$

$$= \underline{e}_{k+1} - \sum_{j=1}^{k} \gamma_{k+1,j} \underline{s}_j = \underline{s}_{k+1}$$

Az (1.26), (1.27) kifejezés érvényességét tehát beláttuk, feltéve az (1.32) kifejezés igaz voltát, amelyet a következő lemmában látunk be.

1.8 Lemma

Az (1.3o) rekurziv képlet a következő alakban irható fel.

$$(1.33) \quad \underline{x}_{k+1}^{m,j} = \underline{x}_k^{m,k} + \underline{e}_{k+1} - \sum_{\ell=1}^{j} \frac{\underline{e}_{k+1}^T A \underline{q}_\ell}{\underline{e}_\ell^T A \underline{q}_\ell} \underline{q}_\ell \quad (j=1,2,\dots,k)$$

Bizonyitás

(1.16), (1.31) szerint (1.3o) a következőképpen irható

$$(1.34)\quad \underline{x}_{k+1}^{m,j} = \underline{x}_{k+1}^{m,j-1} - \frac{b^T \underline{q}_j}{\underline{q}_j^T A \underline{q}_j} \underline{q}_j - \frac{\underline{x}_{k+1}^{m,j^T} A \underline{q}_j}{\underline{q}_j^T A \underline{q}_j} \underline{q}_j =$$

$$= \underline{x}_{k+1}^{m,j-1} + a_j \underline{q}_j - \frac{\underline{x}_{k+1}^{m,j^T} A \underline{q}_j}{\underline{e}_j^T A \underline{q}_j} \underline{q}_j \qquad (j=1,2,\dots,k)$$

A lemmát az (1.34)-re vonatkozó teljes indukcióval látjuk be.

Első lépésként (1.33)-at látjuk be $j=1$ -re. (1.34) alapján a következőket irhatjuk

$$\underline{x}_{K+1}^{m,1} = \underline{x}_{k+1}^{m,o} + a_1 \underline{q}_1 - \frac{\underline{x}_{k+1}^{m,o^T} A \underline{q}_1}{\underline{e}_1^T A \underline{q}_1} \underline{q}_1$$

Felhasználva az $\underline{x}_{k+1}^{m,o}$ -ra vonatkozó (1.29) kifejezést

$$(1.35)\quad \underline{x}_{k+1}^{m,1} = \underline{x}_{k}^{m,k} + \underline{e}_{k+1} + a_1 \underline{q}_1 - \frac{\underline{x}_{k}^{m,k^T} A \underline{q}_1}{\underline{e}_1^T A \underline{q}_1} \underline{q}_1 =$$

$$= \frac{\underline{e}_{k+1}^T A \underline{q}_1}{\underline{e}_1^T A \underline{q}_1} \underline{q}_1$$

Az (1.25) alapján viszont a $\underline{q}_i$, $i=1,2,\dots,k$ irányok A-konjugáltsága miatt

$$(1.36)\quad \frac{\underline{x}_{k}^{m,k^T} A \underline{q}_1}{\underline{e}_1^T A \underline{q}_1} \underline{q}_1 = a_1 \underline{q}_1$$

Ennek alapján (1.35) valóban (1.33) alaku, ugyanis

$$\underline{x}_{k+1}^{m,1} = \underline{x}_{k}^{m,k} + \underline{e}_{k+1} - \frac{\underline{e}_{k+1}^T A \underline{q}_1}{\underline{e}_1^T A \underline{q}_1} \underline{q}_1 .$$

Tegyük fel, hogy (1.33) $j=j_o < k-1$ -ig igaz. Belátjuk, hogy (1.33) $j=j_o+1$-re is igaz. Ennek érdekében irjuk fel (1.34)-et a $j=j_o+1$ esetben.

$$(1.37)\quad \underline{x}_{k+1}^{m,j_o+1} = \underline{x}_{k+1}^{m,j_o} + a_{j_o+1}\,\underline{q}_{j_o+1} - \frac{\underline{x}_{k+1}^{m,j_o^T} A\, \underline{q}_{j_o+1}}{\underline{e}_{j_o+1}^T A\, \underline{q}_{j_o+1}}\,\underline{q}_{j_o+1}$$

Az indukciós feltevés alapján (1.37)-re a következőt kapjuk

$$(1.38)\quad \underline{x}_{k+1}^{m,j_o+1} = \underline{x}_k^{m,k} + \underline{e}_{k+1} - \sum_{\ell=1}^{j_o} \frac{\underline{e}_{k+1}^T A\, \underline{q}_\ell}{\underline{e}_\ell^T A\, \underline{q}_\ell}\,\underline{q}_\ell + a_{j_o+1}\,\underline{q}_{j_o+1} -$$

$$- \frac{\underline{x}_k^{m,k^T} A\, \underline{q}_{j_o+1}}{\underline{e}_{j_o+1}^T A\, \underline{q}_{j_o+1}}\,\underline{q}_{j_o+1} - \frac{\underline{e}_{k+1}^T A\, \underline{q}_{j_o+1}}{\underline{e}_{j_o+1}^T A\, \underline{q}_{j_o+1}}\,\underline{q}_{j_o+1} -$$

$$- \sum_{\ell=1}^{j_o} \frac{\underline{e}_{k+1}^T A\, \underline{q}_\ell}{\underline{e}_\ell^T A\, \underline{q}_\ell} \cdot \frac{\underline{q}_\ell^T A\, \underline{q}_{j_o+1}}{\underline{e}_{j_o+1}^T A\, \underline{q}_{j_o+1}}\,\underline{q}_{j_o+1}\,.$$

Az (1.38) kifejezés utolsó tagja 0, minthogy

$$\underline{q}_\ell^T A\, \underline{q}_{j_o+1} = 0 \qquad (\ell=1,2,\dots,j_o)$$

Az (1.38) kifejezés jobboldalának 5. tagja (1.36)-hoz hasonlóan

$$\frac{\underline{x}^{m,k^T} A\, \underline{q}_{j_o+1}}{\underline{e}_{j_o+1}^T A\, \underline{q}_{j_o+1}}\,\underline{q}_{j_o+1} = a_{j_o+1}\,\underline{q}_{j_o+1}$$

Végül pedig (1.38) jobboldalának 3. és 6. tagját összevonva

$$\underline{x}_{k+1}^{m,j_o+1} = \underline{x}_k^{m,k} + \underline{e}_{k+1} - \sum_{\ell=1}^{j_o+1} \frac{\underline{e}_{k+1}^T A\, \underline{q}_\ell}{\underline{e}_\ell^T A\, \underline{q}_\ell}\,\underline{q}_\ell\,.$$

Az 1.8 lemma segitségével tehát az (1.26) és (1.27) kifejezéseket beláttuk.

Az (1.28) kifejezés belátásához az (1.11) kifejezést használjuk fel. (1.2o) alapján $\underline{x}_{k+1}^{m,k+1}$ -re a következőt kapjuk

$$(1.39)\quad \underline{x}_{k+1}^{m,k+1} = \underline{x}_{k+1}^{m,k} - \frac{\underline{b}^T \underline{g}_{k+1}}{\underline{g}_{k+1}^T A \underline{g}_{k+1}} \underline{g}_{k+1} - \frac{\underline{x}^{m,k}A \underline{g}_{k+1}}{\underline{g}_{k+1}^T A \underline{g}_{k+1}} \underline{g}_{k+1} .$$

Az (1.1o) egyenlőségből viszont

$$\underline{x}_{k+1}^{m,k} = \underline{g}_{k+1} + \underline{x}_k^{m,k}$$

Az $\underline{x}_k^{m,k}$ -ra vonatkozó (1.25) indukciós feltevést az

$$a_{k+1} = - \frac{\underline{b}^T \underline{g}_{k+1}}{\underline{g}_{k+1}^T A \underline{g}_{k+1}}$$

jelölést alkalmazva kapjuk, hogy

$$\underline{x}_{k+1}^{m,k+1} = \underline{g}_{k+1} + \underline{x}_k^{m,k} + \underline{a}_{k+1} \underline{g}_{k+1} - \underline{g}_{k+1} = \sum_{i=1}^{k+1} a_i \underline{g}_i =$$

$$= \sum_{i=1}^{k+1} a_i \underline{s}_i ,$$

amivel az indukciós lépést teljes egészében beláttuk. Bebizonyitottuk tehát a következő tételt.

<u>1.9 Tétel</u>

A Fox-Wilkinson módszer ekvivalens az (1.5)-(1.11) által definiált módszerrel kvadratikus függvények minimalizálásának esetén.

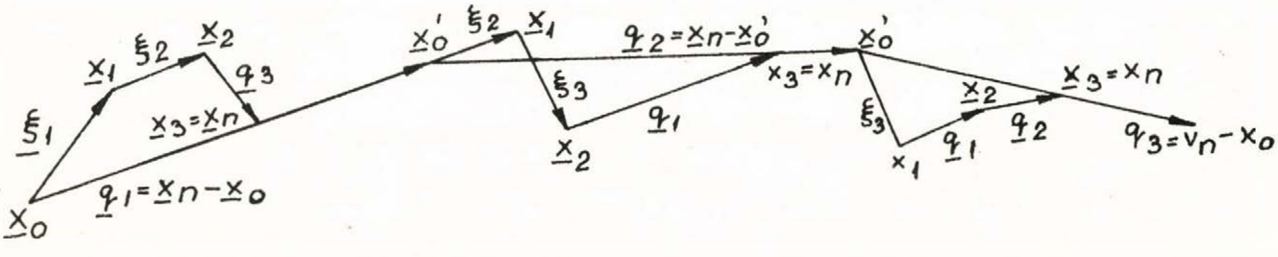
<u>Powell módszer</u>

Legyen $\underline{\xi}_i \neq \underline{O}$, $\underline{\xi}_i \in R^n$, $i=1,2,\ldots,n$, R^n -beli bázis és legyen $\underline{x}_o \in R^n$ tetszőleges.

$$\left.\begin{array}{lll}
(1.4o) & f(\underline{x}_{i-1} + \alpha_i \underline{\xi}_i) = & \\
 & = \min\limits_{\alpha \in R^1} f(\underline{x}_{i-1} + \alpha \underline{\xi}_i) & (i=1,2,\dots,n) \\
(1.41) & \underline{x}_i = \underline{x}_{i-1} + \alpha_i \underline{\xi}_i & \\
(1.42) & \underline{\xi}_i = \underline{\xi}_{i+1} & (i=1,2,\dots,n-1) \\
(1.43) & \underline{\xi}_n = \underline{x}_n - \underline{x}_o = \underline{q}_j & \\
(1.44) & f(\underline{x}_n + \beta_n(\underline{x}_n - \underline{x}_o)) = & \\
 & = \min\limits_{\beta \in R^1} f(\underline{x}_n + \beta(\underline{x}_n - \underline{x}_o)) & \\
(1.45) & \underline{x}_o' = \underline{x}_o + \beta_n(\underline{x}_n - \underline{x}_o) & \\
(1.46) & \underline{x}_o = \underline{x}_o' &
\end{array}\right\} \quad (j=1,2,\dots,n)$$

A következő ábra szemlélteti a Powell módszert, amelyen feltüntettük az iterációs lépéseket is.

1. iterációs lépés 2. iterációs lépés 3. iterációs lépés



A $\underline{q}_1$, $\underline{q}_2$, $\underline{q}_3$ irányok az A-konjugált irányok.

Állapitsuk meg a Powell módszer K_p költségfüggvényének értékét. Legyen a vonalmenti minimalizálásnál kiszámitott függvényértékek száma $\ell_i(\varepsilon) = s$, $i=1,2,\dots,n$ adott $\varepsilon > 0$ esetén. (1.4o)-(1.46)-ból következik, hogy iterációs lépésenként (n+1) vonalmenti minimalizálásra van szükség. Minthogy a Powell módszerre $\omega_i(\varepsilon) = 0$, $i=1,2,\dots,n$ a $K_p(\ell_i\ \omega_i)$ költségfüggvényre kapjuk:

$$(1.47)\quad K_p(\ell_i, \omega_i) = \sum_{i=}^{n} (\ell_i(\varepsilon) + n.\omega_i(\varepsilon)) = \sum_{i=1}^{n} (n+1)s = n(n+1)s$$

azaz

$$(1.48)\qquad V_p = n(n+1)$$

A Powell módszer nem minden kiindulópont esetén kvadratikus befejezésü (lásd [44]). Az alábbiakban leirjuk a Powell módszer Zangwill-től eredő módositását is, amely már kvadratikus befejezésü ([44]).

Legyenek $\underline{e}_i$, $i=1,2,\dots,n$ R^n egység koordináta irányai és $\underline{\xi}_i^1$, $i=1,2,\dots,n$ R^n -beli normált bázis. Legyen $\underline{x}_n^o \in R^n$ tetszőleges induló pont.

$$\underline{x}_{n+1}^o = \underline{x}_n^o + \alpha_n^o\ \underline{\xi}_n^1$$

$$f(\underline{x}_{n+1}^o) = \min_{\alpha \in R^1} f(\underline{x}_n^o + \alpha \underline{\xi}_n^1)$$

$t=1$, $k=1$

1./ Számitsuk ki $\min_{\alpha \in R^1} f(\underline{x}_{n+1}^{k-1} + \alpha\ \underline{e}_t)$ -t, legyen

$$t = \begin{cases} t+1 & \text{ha } 1 \le t < n \\ 1 & \text{ha } t = n \end{cases}$$

ha $\alpha \neq 0$ akkor $\underline{x}_o^k = \underline{x}_{n+1}^{k-1} + \alpha\ \underline{e}_t$

ha $\alpha = 0$ ugrás 1-re. Ha ezt a lépést n-szer egymásután kell elvégezni, akkor $\underline{x}_{n+1}^{k-1}$ minimum pontja $f(\underline{x})$-nek.

2./ $r=1,2,\ldots,n$ -ig számitsuk ki a következőket

$$\underline{x}_r^k = \underline{x}_{r-1}^k + \alpha_r^k \underline{\xi}_r^k$$

ahol α_r^k a következő minimalizálással van meghatározva

$$f(\underline{x}_r^k) = \min_{\alpha \in R^1} f(\underline{x}_{r-1}^k + \alpha \underline{\xi}_r^k)$$

$$\underline{\xi}_{m+1}^k = \frac{\underline{x}_n^k - \underline{x}_{n+1}^{k-1}}{||\underline{x}_n^k - \underline{x}_{n+1}^{k-1}||}$$

$$\underline{x}_{n+1}^k = \underline{x}_n^k + \alpha_{n+1}^k \underline{\xi}_{n+1}^k$$

ahol az α_{n+1}^k paramétert a

$$f(\underline{x}_{n+1}^k) = \min_{\alpha \in R^1} f(\underline{x}_n^k + \alpha_{n+1}^k \underline{\xi}_{n+1}^k)$$

minimalizálás definiálja.

$$\underline{\xi}_r^{k+1} = \underline{\xi}_{r+1}^k \qquad r=1,2,\ldots,n$$

$k = k+1$ ugrás 1-re.

Minthogy a fenti extremális eset csak igen ritkán fordul elő, azért a sokkal müveletigényesebb Zangwill módszert a gyakorlatban nem igen használják. A feladatok többségére az eredeti Powell módszer eredményesebb. Ezért a III. fejezetben a Powell módszert vizsgáljuk.

II. fejezet

Quasi-Newton módszerek

Ebben a fejezetben adunk egy kétparaméteres általános iterációs sémát, amelyről bebizonyitjuk, hogy tetszőleges paraméterválasztás mellett kvadratikus befejezésü. Ebből az általános iterációs sémából vezetjük le az eddig ismert Quasi-Newton módszereket, megfelelő paraméterek megválasztásával.

A fejezetben ismertetendő uj, kvadratikus befejezésü általános iterációs séma, amely tehát a Quasi-Newton módszerek egy osztályát határozza meg tartalmazza a szimmetrikus és nem szimmetrikus módszereket egyaránt. Tulajdonképpen ezzel az általános sémával a Quasi-Newton módszerek egy osztályát határoztuk meg, amely olyan, hogy minden tagja egy pozitiv definit mátrixszal rendelkező kvadratikus alak minimumát legfeljebb n iterációs lépésben (ahol egy iterációs lépésen egy konjugált irány meghatározását értjük) megadja, és az n-edik lépésben a kvadratikus alak Hesse mátrixának inverzét határozza meg.

Mielőtt az uj általános iterációs sémát megadjuk, és annak az 1.2. definició szerinti kvadratikus befejezését bebizonyitjuk, tekintsük az

$$(2.1) \qquad f(\underline{x}) = \frac{1}{2}\,\underline{x}^T A\,\underline{x} + \underline{b}^T\,\underline{x} + c \qquad \underline{x} \in R^n\,,$$

(A n x n-es szimmetrikus pozitiv definit mátrix)

kvadratikus alakot, amelynek gradiens vektora az $\underline{x}$ helyen

$$\underline{g}(\underline{x}) = \text{grad}\ f(\underline{x}) = A\,\underline{x} + \underline{b}$$

Az $f(\underline{x})$ kvadratikus alak minimumhelyét az

(2.2) $A\,\underline{x} = -\,\underline{b}$

lineáris egyenletrendszer megoldása adja meg.

Ez azt jelenti, hogy ha $\underline{g}(\underline{x}) = A\,\underline{x} + \underline{b}$ irány mentén minimalizáljuk az $f(\underline{x})$ kvadratikus alakot, akkor egy lépésben megkapjuk a minimumhelyet. Minthogy az A mátrix inverzére van szükségünk, (2.2)-ből ugyanis

$$\underline{x} = -A^{-1}\,\underline{b}$$

A^{-1} -et kell meghatároznunk. Ehhez segitségünkre van a következő lemma.

(2.1) Lemma

Ha a $\underline{q}_i \neq \underline{0}$ $i=1,2,\dots,n$ irányok A-konjugáltak az n dimenziós térben akkor A^{-1} előáll a következő alakban

$$(2.3) \qquad Z = \sum_{j=1}^{n} \frac{\underline{q}_j\,\underline{q}_j^T}{\underline{q}_j^T\,A\,\underline{q}_j}$$

Bizonyitás

Képezzük az $(I - Z\,A)\,\underline{q}_i$ vektorokat minden $i=1,2,\dots,n$ -re. Kihasználva A pozitiv definitségét és a $\underline{q}_i$ -k A-konjugáltságát, kapjuk:

$$(I - Z\,A)\,\underline{q}_i = \underline{q}_i - \sum_{j=1}^{n} \frac{\underline{q}_j\,\underline{q}_j^T\,A\,\underline{q}_i}{\underline{q}_j^T\,A\,\underline{q}_j} = \underline{q}_i - \underline{q}_i = \underline{0}$$

minden $i=1,2,\dots,n$ esetén, ami viszont azt jelenti, hogy

$$I - Z\,A \equiv 0$$

ahonnan

$$Z = A^{-1}\,.$$

Amennyiben tehát ismerjük a $\underline{q}_1, \underline{q}_2, \dots, \underline{q}_n$ A-konjugált irányokat, a 2.1 lemma alapján eljuthatunk A^{-1} meghatározásához.

Össze kell kapcsolnunk tehát a függvényminimalizáló eljárásokat az A-konjugált irányok meghatározásával. Ehhez fel kell használnunk a Newton módszert, amelyben a Jacobi mátrix a mi esetünkben a kvadratikus alak Hesse mátrixa lesz. Ennek inverzét, a konjugált vektorok meghatározásával párhuzamosan, tehát n iterációs lépésben határozzuk meg. Az igy kapott módszerek kvadratikus befejezésüek. Nem kvadratikus függvények esetére alkalmazva n lépés után nem állunk meg, hanem tovább folytatjuk az eljárást, egész addig, amig az általunk előre megadott kilépési feltételek nem teljesülnek.

Legyen

(2.4) $\quad \underline{q}_i = - H_i^T \underline{g}_i \qquad (i=1,2,\dots,n)$

(2.5) $\quad \underline{x}_{i+1} = \underline{x}_i + \alpha_i \underline{q}_i \qquad (i=1,2,\dots,n)$

ahol H_i n x n -es pozitiv definit mátrix (n a változók száma) $\underline{g}_i$ az $f(\underline{x})$, $\underline{x} \in R^n$ minimalizálandó függvény gradiens vektora az $\underline{x}_i$ helyen, α_i skalár. A $\underline{q}_i$ vektorok lesznek a kvadratikus alak A-konjugált irányai.

Válasszuk meg α_i -t ugy (egyváltozós minimalizálás), hogy a $\underline{q}_i$ irány mentén az $\underline{x}_{i+1}$ hely az $f(\underline{x})$ minimumhelye legyen.
Minthogy a H_i mátrix pozitiv definitségét megköveteltük, a $\underline{q}_i$ irány mindenképpen a függvény csökkenésének irányába mutat (ugyanis nem tudja a gradiens vektort 90^o-kal, vagy többel elforgatni), $\alpha_i \geq 0$ mindig választható. (Ez az észrevétel tulajdonképpen programozás technikai szempontból érdekes).

Teljesüljön a következő feltétel:

(2.6) $\quad H_{j+1} \underline{y}_i = \alpha_i \underline{q}_i \qquad (0 \leq j \leq n)\ ,\ (1 \leq i \leq j)$

ahol

(2.7) $\quad \underline{y}_i = \underline{g}_{i+1} - \underline{g}_i \qquad (i=1,2,\dots)$

A (2.6) feltétel és a (2.7) definició jelentése a következő: a (2.1), (2.5) kifejezést felhasználva a (2.7)-re kapjuk

(2.8) $\quad \underline{y}_i = A\,\alpha_i\,\underline{q}_i$

azaz a (2.3) kifejezésben szereplő nevezők értékét az A felhasználása nélkül kiszámithatjuk (2.7) segitségével, amennyiben a $\underline{q}_i$ irányok A-konjugáltak.

A (2.6) feltételt azért kell megkövetelnünk, mert ennek n-edik tagja az A mátrix inverzét adja, ugyanis (2.8)-at az $\underline{y}_i$ helyébe beirva a

(2.9) $\quad H_{j+1}\,A\,\alpha_i\,\underline{q}_i = \alpha_i\,\underline{q}_i \qquad (1 \le i \le j,\ 0 \le j \le n)$

egyenlőséget kapjuk, ami az $i = n$ esetében azt jelenti, hogy az $\alpha_i\,\underline{q}_i$ vektorok a $H_{n+1}\,A$ mátrix sajátvektorai az 1 sajátértékkel, azaz $H_{n+1}\,A = I$ ahonnan $H_{n+1} = A^{-1}$ adódik.

A (2.6) feltétel a H_i mátrixsorozatot nem határozza meg egyértelmüen /még akkor sem, ha a pozitiv definitségen kivül a mátrix szimmetricitását is feltesszük/. Ezért a H_i mátrixsorozatra további megszoritásokat is kell tennünk.

Legyen H_1 tetszőleges pozitiv definit szimmetrikus mátrix, és

(2.10) $\quad H_{i+1} = H_i + C_i \qquad (i=1,2,\dots)$

Megszorozva (2.10)-et $\underline{y}_i$ -vel és a (2.6) feltételt felhasználva kapjuk

(2.11) $\quad C_i\,\underline{y}_i = \alpha_i\,\underline{q}_i - H_i\,\underline{y}_i \qquad (i=1,2,\dots)$

Ha $\underline{z}_i \in R^n$, ($\underline{z}_i \neq 0$) olyan vektor, amelyre

$$z_i^T y_i = 1 \qquad (i=1,2,\ldots) \tag{2.12}$$

akkor a

$$C_i = (\alpha_i q_i - H_i y_i) z_i^T \qquad (i=1,2,\ldots) \tag{2.13}$$

egyenletből a (2.11) következik. A (2.13) kifejezés további általánositása a következő

$$C_i = \alpha_i q_i s_i^T - H_i y_i z_i^T \qquad (i=1,2,\ldots) \tag{2.14}$$

amelyben s_i -re hasonló egyenlőséget követelünk meg mint a z_i vektorra:

$$s_i^T y_i = 1 \tag{2.15}$$

Az eddigi meggondolások alapján a következő általános iterációs sémát adhatjuk meg:

Legyen H_1 tetszőleges pozitiv definit mátrix, $x_1 \in R^n$ tetszőleges kiinduló pont, továbbá

$$q_i = - H_i^T g_i \tag{2.16}$$

$$x_{i+1} = x_i + \alpha_i q_i \qquad (i=1,2,\ldots,n) \tag{2.17}$$

$$y_i = g_{i+1} - g_i \tag{2.18}$$

$$H_{i+1} = H_i + \alpha_i q_i s_i^T - H_i y_i z_i^T \tag{2.19}$$

amelyben $\alpha_i \geq 0$ mindig választható és s_i -re, z_i -re a (2.12) és (2.15) feltételek teljesülnek.

Stabilitási definiciót a fenti iterációs sémára a következőképpen adhatunk:

2.2 Definició

Azt mondjuk, hogy a fenti általános iterációs séma stabil, ha a/ H_i pozitiv definit mátrix, bármely i=1,2,..., esetén

b/

$$\underline{s}_i^T \underline{y}_i = 1 \quad \text{és} \quad \underline{z}_i^T \underline{y}_i = 1 \qquad (i=1,2,\ldots) \tag{2.20}$$

A 2.2 stabilitási definició választása természetes, mert egyrészt a (2.16), (2.17) miatt igy $\alpha_i \geq 0$ minden $i=1,2,\ldots,n$ esetben teljesül, másrészt célszerü a (2.1)-ben szereplő A szimmetrikus pozitiv definit mátrix inverzét, amely szintén szimmetrikus pozitiv definit mátrix, egy pozitiv definit mátrix-sorozaton keresztül meghatározni. A (2.20) feltételek a (2.12), (2.15) feltételekből adódnak.

Nem következik a 2.2 definicióból viszont az, hogy csak stabil módszerek lehetnek kvadratikusan konvergensek.

Az $\underline{s}_i$, $\underline{z}_i$ vektorokat a következőképpen határozzuk meg:

Legyen

$$\underline{s}_i = \beta_i \underline{q}_i + \frac{1 - \beta_i \underline{q}_i^T \underline{y}_i}{\underline{y}_i^T H_i \underline{y}_i} H_i^T \underline{y}_i \qquad (i=1,2,\ldots,n) \tag{2.21}$$

$$\underline{z}_i = \frac{\alpha_i \delta_i \underline{q}_i^T \underline{y}_i + 1}{\underline{y}_i^T H_i \underline{y}_i} H_i^T \underline{y}_i - \alpha_i \delta_i \underline{q}_i \qquad (i=1,2,\ldots,n) \tag{2.22}$$

ahol β_i, δ_i tetszőlegesen választható konstansok.

Az $\underline{s}_i$, $\underline{z}_i$ vektorok ilyen megválasztásából azonnal adódik az alábbi következmény:

2.3 <u>Következmény</u>

A 2.2 stabilitási definició (2.20) feltételei tetszőleges β_i, δ_i választás mellett teljesülnek.

Bizonyitás

$$\underline{s}_i^T \underline{y}_i = \beta_i \underline{g}_i^T \underline{y}_i + \frac{1 - \beta_i \underline{g}_i^T \underline{y}_i}{\underline{y}_i^T H_i \underline{y}_i} \underline{y}_i^T H_i \underline{y}_i = 1$$

$$\underline{z}_i^T \underline{y}_i = \frac{\alpha_i \delta_i \underline{g}_i^T \underline{y}_i + 1}{\underline{y}_i^T H_i \underline{y}_i} \underline{y}_i^T H_i \underline{y}_i - \alpha_i \delta_i \underline{g}_i^T \underline{y}_i = 1$$

A (2.21), (2.22) kifejezéseket (2.19)-be helyettesitve kapjuk a következő, kétparaméteres (β_i, δ_i) általános iterációs sémát.

Legyen H_1 tetszőleges pozitiv definit mátrix, $\underline{x}_1 \in R^n$ tetszőleges, továbbá

(2.23) $\underline{g}_i = - H_i^T \underline{g}_i$

(2.24) $\underline{x}_{i+1} = \underline{x}_i + \alpha_i \underline{g}_i$

(2.25) $f(\underline{x}_{i+1}) = \min\limits_{\alpha \in R^1} f(\underline{x}_i + \alpha\, \underline{g}_i)$ $\qquad$ (i=1,2,...,n)

(2.26) $\underline{y}_i = \underline{g}_{i+1} - \underline{g}_i$

(2.27)
$$H_{i+1} = H_i - \frac{\alpha_i \delta_i \underline{g}_i^T \underline{y}_i + 1}{\underline{y}_i^T H_i \underline{y}_i} H_i \underline{y}_i \underline{y}_i^T H_i + $$
$$+ \alpha_i \beta_i \underline{g}_i \underline{g}_i^T + \alpha_i \delta_i H_i \underline{y}_i \underline{g}_i^T +$$
$$+ \alpha_i \frac{1 - \beta_i \underline{g}_i^T \underline{y}_i}{\underline{y}_i^T H_i \underline{y}_i} \underline{g}_i \underline{y}_i^T H_i$$

ahol $\underline{g}_i = \text{grad } f(\underline{x}_i)$ (i=1,2,...,n).

2.4 Tétel

Ha H_1 tetszőleges pozitiv definit mátrix, $\underline{x}_1 \in R^n$ egy tetszőleges kiinduló pont, akkor a (2.23) - (2.27)-ban meghatározott iterációs séma tetszőleges β_i, δ_i $(i=1,2,\dots,n)$ paraméter választás mellett kvadratikus befejezésü.

Bizonyitás

Be kell látnunk, hogy egy (2.1) alaku pozitiv definit Hesse mátrixszal rendelkező kvadratikus alak minimumát legfeljebb n iterációs lépésben megkapjuk. Egy iterációs lépés nyilván a (2.23) - (2.27) képletek egyszeri kiszámitását jelenti.

A bizonyitást teljes indukcióval végezzük el. Feltesszük, hogy

$$H_k A \alpha_i \underline{q}_i = \alpha_i \underline{q}_i \qquad (1 \le i < k) \tag{2.28}$$

$$\underline{q}_j^T A \underline{q}_i = 0 \qquad (1 \le i < j < k) \tag{2.29}$$

és belátjuk, hogy $k + 1 \le n$ -re is igazak a feltételek.

Az indukciós bizonyitás első lépéseként (2.28)-ban legyen $k = 2$. Felhasználva (2.8)-at és (2.27)-et kapjuk:

(2.30)

$$H_2 A \alpha_1 \underline{q}_1 = H_2 \underline{y}_1 = H_1 \underline{y}_1 -$$

$$- \frac{\alpha_1 \delta_1 \underline{q}_1^T \underline{y}_1 + 1}{\underline{y}_1^T H_1 \underline{y}_1} H_1 \underline{y}_1 \underline{y}_1^T H_1 \underline{y}_1 + \alpha_1 \beta_1 \underline{q}_1^T \underline{y}_1 \underline{q}_1 +$$

$$+ \alpha_1 \delta_1 H_1 \underline{y}_1 \underline{q}_1^T \underline{y}_1 + \alpha_1 \frac{1 - \beta_1 \underline{q}_1^T \underline{y}_1}{\underline{y}_1^T H_1 \underline{y}_1} \underline{q}_1 \underline{y}_1^T H_1 \underline{y}_1 =$$

$$= H_1 \underline{y}_1 - \alpha_1 \delta_1 \underline{q}_1^T \underline{y}_1 H_1 \underline{y}_1 - H_1 \underline{y}_1 + \alpha_1 \beta_1 \underline{q}_1 \underline{q}_1^T \underline{y}_1 +$$

$$+ \alpha_1 \delta_1 H_1 \underline{y}_1 \underline{q}_1^T \underline{y}_1 + \alpha_1 \underline{q}_1 - \alpha_1 \beta_1 \underline{q}_1 \underline{q}_1^T \underline{y}_1 = \alpha_1 \underline{q}_1$$

(2.28)-ban első lépésként legyen $k = 3$, (2.8)-at, (2.23)-at és (2.28)-at $k = 2$ -re alkalmazva kapjuk:

$$\underline{g}_2^T A \underline{g}_1 = -\underline{g}_2^T H_2 A \underline{g}_1 = -\underline{g}_2^T H_2 \frac{\underline{y}_1}{\alpha_1} = -\underline{g}_2^T \underline{g}_1 = 0,$$

ugyanis a $\underline{g}_i$ irány mentén (2.25) szerint meghatároztuk az $\underline{x}_2 \in R^n$ minimumhelyet, igy szükségképpen a $\underline{g}_2 = \text{grad } f(\underline{x}_2)$ a $\underline{g}_1$ -re ortogonális.

Az indukciós bizonyitás általános lépéseként először belátjuk (2.28)-at $k + 1$ -re.

Teljesen hasonlóan (2.30)-hoz, a 2-es index helyébe $k + 1$-et, az 1-es indexek helyébe pedig k-t irva kapjuk

$$H_{k+1} A \alpha_k \underline{g}_k = H_{k+1} \underline{y}_k = \alpha_k \underline{g}_k .$$

Feltehetjük tehát /k helyébe $k+1$-et irva/, hogy

$$1 \le i < k$$

A (2.8), (2.27) és a (2.28) indukciós feltevést ujből felhasználva (2.28) a következőképpen alakul

$$(2.31) \quad H_{k+1} A \alpha_i \underline{g}_i = H_k A \alpha_i \underline{g}_i -$$

$$- \frac{\alpha_k \delta_k \underline{g}_k^T \underline{y}_k + 1}{\underline{y}_k^T H_k \underline{y}_k} H_k \underline{y}_k \underline{y}_k^T H_k A \alpha_i \underline{g}_i +$$

$$+ \alpha_k \beta_k \underline{g}_k \underline{g}_k^T A \alpha_i \underline{g}_i + \alpha_k \delta_k H_k \underline{y}_k \underline{g}_k^T A \alpha_i \underline{g}_i +$$

$$+ \alpha_k \frac{1 - \beta_k \underline{g}_k^T \underline{y}_k}{\underline{y}_k^T H_k \underline{y}_k} \underline{g}_k \underline{y}_k^T H_k A \alpha_i \underline{g}_i = \alpha_i \underline{g}_i \quad (1 \le i < k)$$

ugyanis (2.31) jobb oldalának első tagja az indukciós feltevés miatt $\alpha_i \underline{g}_i$. A 2. és 5. tagra vonatkozóan pedig vizsgáljuk az $\underline{y}_k^T H_k A \alpha_i \underline{g}_i$ mennyiségeket. A (2.8) és a (2.27) indukciós feltevés miatt

(2.32) $\quad \underline{y}_k^T H_k A \alpha_i \underline{g}_i = \underline{y}_k^T \alpha_i \underline{g}_i = \alpha_k \underline{g}_k^T A \alpha_i \underline{g}_i$

amelyben felhasználtuk az A Hesse mátrix szimmetricitását.

Azt kaptuk tehát, hogy (2.31) jobb oldalának első tagját kivéve, minden további tagban megjelenik a

(2.33) $\quad \underline{g}_k^T A \underline{g}_i \qquad (1 \le i < k)$

mennyiség. $\underline{g}_k^T A \underline{g}_i = 0$, ha belátjuk a (2.29) feltevést k+1 -re.

Ekkor ugyanis (2.29)-ben $j = k$ -t irhatunk.

Ezt a következőképpen láthatjuk be (2.24) alapján

$$\underline{x}_k = \underline{x}_{k-1} + \alpha_{k-1} \underline{g}_{k-1} = \underline{x}_{i+1} + \sum_{\ell=i+1}^{k-1} \alpha_\ell \underline{g}_\ell \quad (i+1 < k)\text{-ra.}$$

Megszorozva balról az A mátrixszal

$$A \underline{x}_k = A \underline{x}_{i+1} + \sum_{\ell=i+1}^{k-1} \alpha_\ell A \underline{g}_\ell$$

mindkét oldalhoz a $\underline{b}$ vektort hozzáadva

$$\underline{g}_k = \underline{g}_{i+1} + \sum_{\ell=i+1}^{k-1} \alpha_\ell A \underline{g}_\ell .$$

A fentiek alapján

$$\underline{g}_k^T \underline{g}_i = \underline{g}_{i+1} \underline{g}_i + \sum_{\ell=i+1}^{k-1} \alpha_\ell \underline{g}_\ell^T A \underline{g}_i$$

tetszőleges $i < k$ -ra

de az indukciós feltevés miatt

$$\sum_{\ell=i+1}^{k-1} \alpha_\ell \underline{g}_\ell^T A \underline{g}_i = 0, \qquad \underline{g}_{i+1}^T \underline{g}_i = 0$$

hiszen (2.25) miatt $\underline{g}_{i+1}$ ortogonális $\underline{g}_i$ -re.

Tehát

(2.34) $\quad \underline{g}_k^T \underline{g}_i = 0$ tetszőleges $i < k$ -ra.

Mármost (2.33) a következőképpen irható:

$$\underline{g}_k^T A \underline{g}_i = - \underline{g}_k^T H_k A \underline{g}_i = - \underline{g}_k^T \underline{g}_i = 0 \qquad (1 \le i < k)$$

(2.34) miatt, tehát (2.31) helyességét és a (2.29) indukciós feltételt is beláttuk.

Az n-edik lépés után a $\underline{g}_i \neq \underline{0}$ $(i=1,2,\dots,n)$ vektorok kifeszitik R^n -t. /Amennyiben valamelyik $\underline{g}_j = \underline{0}$ fennállna, ugy a megoldást már a j-edik lépésben megkapnánk/. Ennek alapján (2.28)-ra a következőt mondhatjuk

$$(2.35) \qquad H_{n+1} A \underline{w}_j = \underline{w}_j \qquad (j=1,2,\dots,n)$$

ahol $\underline{w}_j = \alpha_j \underline{g}_j$ $(j=1,2,\dots,n)$. Viszont (2.35) azt jelenti, hogy a $\underline{w}_j$, $(j=1,2,\dots,n)$ vektorok sajátvektorai a $H_{n+1} A$ mátrix'nak az 1 sajátértékkel.

Ebből következik

$$H_{n+1} A = I \qquad \text{és innen}$$

$$H_{n+1} = A^{-1} \qquad \text{minthogy } A \text{ pozitiv definit.}$$

Az 1.5 lemma miatt viszont $\underline{x}_{n+1}$ az $f(\underline{x})$ kvadratikus alak minimumhelye R^n -ben.

Ezzel a 2.4 tételt bebizonyitottuk.

A Quasi-Newton módszerek költségfüggvénye (K_Q) tehát

$$K_Q = \sum_{i=1}^{n} (1+ n) = n(n + 1)$$

A továbbiakban megmutatjuk azt, hogy az eddig ismert Quasi--Newton módszerek a (2.23)-(2.27) sémából hogyan nyerhetők.

A/ Szimmetrikus Quasi-Newton módszerek.

1/ Fletcher-Powell-Davidon algoritmus [15]

Válasszuk meg a β_i, δ_i paramétereket 2.27 -ben a következő módon

$$\beta_i = \frac{\alpha_i}{\alpha_i \underline{q}_i^T \underline{y}_i}, \qquad \delta_i = 0 \qquad (i=1,2,\ldots,n)$$

Látható, hogy a H_i $(i=1,2,\ldots,n)$ mátrixsorozat szimmetrikus, amennyiben a H_1 is az. A Fletcher-Powell-Davidon módszer ennek megfelelően a következő:

$$\underline{q}_i = - H_i \underline{g}_i \tag{2.36}$$

$$\underline{x}_{i+1} = \underline{x}_i + \alpha_i \underline{q}_i \tag{2.37}$$

$$(i=1,2,\ldots,n)$$

$$f(\underline{x}_{i+1}) = \min_{\alpha \in R^1} f(\underline{x}_i + \alpha_i \underline{q}_i) \tag{2.38}$$

$$\underline{y}_i = \underline{g}_{i+1} - \underline{g}_i \tag{2.39}$$

$$H_{i+1} = H_i + \frac{\alpha_i \underline{q}_i \alpha_i \underline{q}_i^T}{\alpha_i \underline{q}_i^T \underline{y}_i} - \frac{H_i \underline{y}_i \underline{y}_i^T H_i}{\underline{y}_i^T H_i \underline{y}_i} \tag{2.40}$$

2/ Broyden szimmetrikus Quasi-Newton osztálya [7]

Válasszuk meg a β_i, δ_i paramétereket (2.27)-ben a következő módon:

$$\beta_i = \frac{1 + \vartheta_i \underline{y}_i^T H_i \underline{y}_i}{\underline{q}_i^T \underline{y}_i}, \qquad \delta_i = - \vartheta_i \qquad (i=1,2,\ldots,n) \tag{2.41}$$

ahol ϑ_i $(i=1,2,\ldots,n)$ tetszőleges konstans, akkor a Broyden osztályt kapjuk. Ebből következően Broyden szimmetrikus Quasi-Newton osztálya, az általános iterációs séma által meghatározott módszerosztály része. Behelyettesitve β_i, δ_i -t (2.27)-be kapjuk a következőt:

Az első négy egyenlet egybeesik (2.36)-(2.39)-cel majd

$$H_{i+1} = H_i - \frac{1 - \vartheta_i \alpha_i \underline{q}_i^T \underline{y}_i}{\underline{y}_i^T H_i \underline{y}_i} H_i \underline{y}_i \underline{y}_i^T H_i +$$

$$+ \frac{1 + \vartheta_i \underline{y}_i^T H_i \underline{y}_i}{\underline{q}_i^T \underline{y}_i} \alpha_i \underline{q}_i \underline{q}_i^T - \vartheta_i \alpha_i (H_i \underline{y}_i \underline{q}_i^T + \underline{q}_i \underline{y}_i^T H_i)$$

és

természetesen H_1 -et szimmetrikus pozitiv definitnek kell választani.

2.5 Következmény

A (2.23)-(2.27) által definiált módszerosztályban a szimmetrikus módszerek részosztályát a Broyden osztály teljesen lefedi.

Bizonyitás

Ahhoz, hogy szimmetrikus mátrixsorozatot kapjunk (2.27) alapján a következő egyenletnek kell fennállnia

$$\alpha_i \delta_i = \alpha_i \frac{1 - \beta_i \underline{q}_i^T \underline{y}_i}{\underline{y}_i^T H_i \underline{y}_i} \qquad (i=1,2,\dots,n) \tag{2.42}$$

s ez valóban equivalens a (2.41) feltételekkel.

Megjegyzés.

1. Következésképpen a Fletcher-Powell-Davidon módszer eleme a Broyden osztálynak. Valóban, ha $\vartheta_i = 0$ $(i=1,2,\dots,n)$ akkor éppen a (2.40) kifejezést kapjuk.

2. Az uj általános iterációs sémának, lényeges tulajdonsága az, hogy az n-edik lépés után

$$H_{n+1} A = I .$$

Amennyiben ezt a feltételt feladjuk, ugy megfogalmazható egy még általánosabb módszerosztály, amelyben a Broyden osztály már nem fedi le az összes szimmetrikus Quasi-Newton módszert. Lásd Abaffy [1].

B/ Nem szimmetrikus Quasi-Newton módszerek:

Amennyiben a (2.42) egyenlet nem áll fenn β_i és γ_i (i=1,2,...,n között, akkor a H_i (i=1,2,...,n) mátrixsorozat nem lesz szimmetrikus, függetlenül a H_1 induló mátrix megválasztásától. Minthogy ezeknek a módszereknek első 4 egyenlete egybeesik (2.23)-(2.26)-tal, a továbbiakban csak a (2.27) egyenlet megfelelő módositását irjuk le. A következő módszerek ismeretesek.

1/ Mc Cormick módszere [32]

Válasszuk meg a β_i , δ_i (i=1,2,...,n) paramétereket (2.27)-ben a következő módon:

$$\beta_i = \frac{\alpha_i}{\alpha_i \underline{q}_i^T \underline{y}_i}, \qquad \delta_i = - \frac{1}{\alpha_i \underline{q}_i^T \underline{y}_i} \qquad (i=1,2,...,n)$$

Akkor (2.27) helyébe a következő kifejezés lép:

$$H_{i+1} = H_i + \frac{\alpha_i \underline{q}_i \alpha_i \underline{q}_i^T}{\alpha_i \underline{q}_i^T \underline{y}_i} - \frac{H_i \underline{y}_i \alpha_i \underline{q}_i^T}{\alpha_i \underline{q}_i^T \underline{y}_i} \qquad (i=1,2,...,n)$$

2/ Pearson módszere [32]

Legyen $\beta_i = 0$, $\delta_i = 0$ (i=1,2,...,n) (2.27)-ben. Akkor (2.27) a következőképpen alakul:

$$H_{i+1} = H_i + \frac{\alpha_i \underline{q}_i \underline{y}_i^T H_i}{\underline{y}_i^T H_i \underline{y}_i} - \frac{H_i \underline{y}_i \underline{y}_i^T H_i}{\underline{y}_i^T H_i \underline{y}_i} \qquad (i=1,2,..,n)$$

III. fejezet

Gyakorlati eredmények

Ebben a fejezetben az előbbiekben ismertetett módszerek próbafüggvényeken való lefuttatásán /amelyeket az MTA CDC 3300-as gépén végeztünk el/ és az ezzel kapcsolatos eredmények elemzésén kivül egy, a Quasi-Newton módszerek osztályára vonatkozó gyorsitási lehetőséget is ismertetünk.

1. A függvényminimalizálás módszereit általában nem kvadratikus függvények minimumhelyének és értékének megkeresésére használják. Innen rögtön adódik az /és ezt a próbafüggvényeken való futtatásaink is igazolják/, hogy bár a Quasi-Newton módszerek költségfüggvénye azonos minden módszerre, maguk a módszerek viszont nagyon különbözőképpen viselkednek a nem kvadratikus függvényekre történő alkalmazásuk során. Mostanáig sem ismeretes olyan tétel, vagy módszer, amely egy adott függvényosztályra optimális. Ezért alakultak ki a függvényminimalizálás irodalmában az ugynevezett próbafüggvények, amelyeken minden uj módszert "illik" kipróbálni. Az általunk adott költségfüggvény az első próbálkozás az egyes módszerek "jóságának" meghatározására. A költségfüggvény alapján /amint ezt a későbbiekben látni fogjuk/ a Quasi-Newton módszerek egy bizonyos "jósági" sorrendbe állithatók. Ez a kijelentés természetesen csak statisztikai jellegü. A kapott sorrendet egyébként nem csak az 5 próbafüggvényen történt futtatásaink igazolják, hanem egyéb, itt hely hiánya miatt nem közölhető futtatásaink, valamint a nemzetközi irodalomban közölt vizsgálatok is /pl. Himmelblau könyvében [20]/.

Ugy tünik, hogy a módszerek közti eltérések elsősorban a számitások során fellépő, és az egyes módszereknél különbözőképpen halmozódó hibákból erednek.

A fenti észrevételek igazak az I. fejezetben ismertetett Smith és Powell módszerre is. A költségfüggvény alapján a Smith módszer jobb, mint a Powell módszer, hiszen a vonalmenti minimalizálások száma /és általában ez a müveletigényes/ fele a Powell módszerénél, azonban egy hátránya azonnal láthatóvá válik. Ez pedig az, hogy amig az első $\underline{q}_1$ konjugált irány mentén n vonalmenti minimalizálást végzünk, addig az utolsó $\underline{q}_n$ konjugált irány mentén csupán egyet. A $\underline{q}_1$ konjugált irány kitüntetettsége nem indokolt. Az ebből adódó hátrány kvadratikus függvények esetén is érezteti hatását a számitások során fellépő hibák miatt.

A futtatásokat a következő próbafüggvényekre végeztük el:

(A) $f(\underline{x}) = 100(x_2 - x_1^2)^2 + (1 - x_1)^2$, induló pont (-1.2,1)
a minimum helye (1,1), a minimum értéke 0

(B) $f(\underline{x}) = (x_2 - x_1^2)^2 + (1 - x_1)^2$, induló pont (-1.2,1)
a minimum helye (1,1), a minimum értéke 0

(C) $f(\underline{x}) = (x_1 + 10x_2)^2 + 5(x_3 - x_4)^2 + (x_2 - 2x_3)^4 +$
$+ 10(x_1 - x_4)^4$, induló pont (-3,-1,0,1), a minimum helye (0,0,0,0) a minimum értéke 0

(D) $f(\underline{x}) = (\exp(x_1) - x_2)^4 + 100(x_2 - x_3)^6 + \tan^4(x_3 - x_4) +$
$+ x_1^8 + (x_4 - 1)^2$ induló pont (1,2,2,2), a minimum helye $(0,1,1,1,\pm n\pi)$ a minimum értéke 0

(E) $f(\underline{x}) = x_1(2x_1 - x_3 + 1) + x_2(x_2 - 3) + x_3(2x_3 + x_4 + 1) +$
$+ x_4(x_4 - 1)$ induló pont (20,20,20,20), a minimum helye
$(^2/13,\ ^3/2,\ -^5/13,\ ^9/13)$ a minimum értéke - 2.86538...

Az alábbi 1. táblázatban közöljük a Smith és a Powell módszerre kapott számitógépes eredményeket. A Smith módszer esetén a $\beta_i \neq 0$, $i=1,2,\ldots$ tetszőleges konstans meghatározására a következő feltételt szabtuk ki. Válasszuk a β_i -t ($\beta_i \neq 0$) ugy, hogy az

$$f(\underline{x}_i') < f(\underline{x}_{i-1}^{m,i-1} + \beta_i \underline{e}_i)$$

feltétel teljesüljön minden i-re.
A költségfüggvény rovatba a tényleges függvényértékek kiszámitásának a számát irtuk be. Jelen esetben gradiens vektort ugyanis nem kell kiszámitanunk.

Vonalmenti minimalizálásra az aranymetszési módszert használtuk. Ennek kilépési feltételeként a következő két feltételt állitottuk logikai "és" kapcsolatba

(3.1) $$\left|\frac{f(\underline{x}_k) - f(\underline{x}_{k+1})}{f(\underline{x}_k)}\right| < 10^{-5}$$

(3.2) $$\| \underline{x}_k - \underline{x}_{k+1} \|_{\ell_2} < 10^{-5}$$

valamilyen k-ra.

A Smith módszer és a Powell módszer kilépési feltételeként pedig az alábbi két feltétel egyszerre való teljesülését határoztuk meg:

$$\left| \frac{f(\underline{x}_k^{min}) - f(\underline{x}_{k+1}^{min})}{f(\underline{x}_k^{min})} \right| < 10^{-5} \tag{3.3}$$

két egymásutáni minimum értékre, illetve

$$\left\| \underline{x}_k^{min} - \underline{x}_{k+1}^{min} \right\|_{\ell_2} < 10^{-5} \tag{3.4}$$

két egymásutáni minimumhelyre.

Ezenkivül az egyes számitásokra a következő korlátozásokat vezettük még be:

a/ leállás, amennyiben az iterációk száma meghaladja az 1500-at;

b/ leállás, amennyiben a feladat számára felhasznált gépidő több 3 percnél.

Az alábbi 1. táblázat mutatja a számitási eredményeket. A kapott tényleges költségfüggvények alapján megállapitható, hogy a Powell módszer jobb a Smith módszernél. A módszer rovatban PO-val Powell, S-sel Smith módszerét jelöljük.

Próba fv.	Módszer	Induló pont	Min. érték	Minimum hely x_1 x_2	x_3 x_4	Költség fv.	Futt.idő /sec/	Iteráció szám
A	PO	(-1,2,1)	$0.3506.10^{-10}$	1.000003 1.000006		532	0.789	25
	S		$0.4622.10^{-3}$	0.978501 0.957489		1075	0.863	36
B	PO	(-1,2,1)	$0.1744.10^{-9}$	0.999986 0.999973		148	0.378	3
	S		$0.5460.10^{-11}$	0.999997 0.999995		294	0.251	6
C	PO	(-3,-,1., 0,1)	$0.2166.10^{-25}$	$0.303733.10^{-6}$ $-0.303733.10^{-6}$	$0.104447.10^{-6}$ $-0.104447.10^{-6}$	2669	5.322	107
	S		$0.4372.10^{-6}$	$0.216767.10^{-2}$ $-0.216258.10^{-2}$	$0.896170.10^{-2}$ $0.901644.10^{-2}$	82496	159.862	1500
D	PO	(1,2, 2,2)	$0.3457.10^{-20}$	$0.915312.10^{-4}$ 1.000092	0.999998 1.000000	2065	142.345	76
	S		$0.7021.10^{-8}$	$-0.498368.10^{-4}$ 0.999955	0.999955 0.999916	32858	133.420	498
E	PO	(20,20, 20,20)	-2.8654	0.153845 1.500002	-0.384612 0.692303	825	2.548	25
	S		-2.8654	0.153845 1.500000	-0.384616 0.692306	1112	1.404	18

1. táblázat

A II. fejezetben ismertetett módszereket ugyancsak az (A) - (E) próbafüggvényekkel futtattuk le. A futtatási eredményeket a 2. táblázat tartalmazza. A módszerek rovatban F - P -vel a Fletcher-Powell-Davidon módszert, P-vel a Pearson módszert, C-vel pedig az Mc Cormick módszert jelöltük. A vonalmenti minimalizáláshoz ismét az aranymetszési módszert használtuk a (3.1) és a (3.2) kilépési feltételekkel. Az egyes módszerek kilépési feltételeként a programba a (3.3) és a (3.4) feltételeken kivül a következő feltételt épitettük be "és" kapcsolattal

$$\| \underline{g}(\underline{x}_{k+1}^{min}) \|_{\ell_2} \le 10^{-5} \tag{3.5}$$

A tényleges költségfüggvényértékbe most a gradiens irány kiszámitását is beleszámitottuk.

Próba fv.	Módszer	Induló pont	Min. érték	Minimum hely x_1 x_2	x_3 x_4	Költség fv.	Futt.idő /sec/	Iteráció szám
A	F-P	(-1.2,1)	$0.6231.10^{-15}$	1.000000 1.000000		146	0.254	12
	P		$0.2076.10^{-14}$	1.000000 1.000000		806	1.149	64
	C		0.19461	1.440885 2.077677		51826	44.482	1500
B	F-P	(-1.2,1)	$0.7428.10^{-18}$	1.000000 1.000000		84	0.125	7
	P		$0.1694.10^{-18}$	1.000000 1.000000		121	0.224	10
	C		0.0140	1.100112 1.273486		83086	12.871	1500
C	F-P	(-3,-1, 0,1)	$0.2235.10^{-17}$	$0.149775.10^{-7}$ $0.149775.10^{-7}$	$0.100661.10^{-7}$ $0.107346.10^{-7}$	3900	8.185	391
	P		$0.4309.10^{-20}$	$0.666475.10^{-5}$ $0.66475.10^{-6}$	$0.342274.10^{-5}$ $0.342274.10^{-5}$	28054	63.444	1500
	C		3.3519	-0.348852 -0.052567	-0.411017 -0.910009	65165	122.907	1500
D	F-P	(1,2, 2,2)	$0.1250.10^{-32}$	$-0.144911-10^{-5}$ 0.999998	1.000000 1.000000	4878	15.218	552
	P		$0.1091.10^{-5}$	0.027984 1.017882	0.994168 0.998969	45075	142.326	1500
	C		0.1491	0.530769 1.480850	1.645115 1.362611	58971	161.783	1500
E	F-P	(20,20, 20,20)	-2.8654	0.153846 1.500000	-0.384615 0.692307	139	0.306	6
	P		-2.8654	0.153846 1.500000	-0.384615 0.692307	162	0.358	9
	C		-2.8654	0.153846 1.500000	-0.384615 0.692307	249	0.428	13

2. táblázat

A táblázatból a költségfüggvény alapján megállapitható, hogy a Fletcher-Powell-Davidon módszer lényegesen "jobb" a többi módszernél.

2. Quasi-Newton módszerek egy gyorsitási lehetősége.

Tegyük fel, hogy a H_i mátrixsorozat, amellyel a Quasi-Newton módszerekben egy kvadratikus alak Hesse mátrixának inverzét közelitjük a pozitiv definitségen kivül szigoruan diagonálisan domináns, azaz

$$|h_{ii}| > \sum_{\substack{j=1 \\ j\neq i}}^{n} |h_{ij}| \quad , \qquad (i=1,2,\dots,n)$$

Ha a diagonális elem valóban sokkal nagyobb mint a sor többi eleme abszolut értékének összege, akkor ez azt jelenti, hogy

$$\sum_{j=1}^{n} |h_{ij}| \sim c\sqrt{\lambda_i} \quad , \qquad (i=1,2,\dots,n)$$

Legyen

$$\max_{1\le i\le n} \sum_{j=1}^{n} |h_{ij}| = M$$

$$\min_{1\le i\le n} \sum_{j=1}^{n} |h_{ij}| = m > 0 \ ,$$

ebből

$$\frac{M}{m} \sim \frac{c\sqrt{\lambda_s}}{c\sqrt{\lambda_k}} = \frac{\sqrt{\lambda_s}}{\sqrt{\lambda_k}} \ ,$$

és s a maximális, k a minimális sajátértékhez tartozó index.

Végezzük el a következő transzformációt:

$$L = \begin{array}{c l} \begin{pmatrix} 1 & 0 & \dots & 0 & 0 & 0 & \dots & 0 \\ 0 & 1 & & 0 & 0 & 0 & & \cdot \\ \cdot & \cdot & & \cdot & \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot & \cdot & \cdot & & \cdot \\ 0 & 0 & \dots & 1 & 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & \frac{M}{m} & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & 0 & 1 & \dots & 0 \\ \cdot & & & & \cdot & & & \cdot \\ \cdot & & & & \cdot & & & \cdot \\ \cdot & & & & \cdot & & & \cdot \\ 0 & 0 & \dots & 0 & 0 & 0 & \dots & 1 \end{pmatrix} & \begin{array}{l} \\ \\ \\ \\ \\ \text{k-adik sor} \\ \\ \\ \\ \\ \\ \end{array} \end{array}$$

$$H_i' = L\, H_i\, L^T \qquad \underline{x}' = L^{-1}\, \underline{x}$$

Szemléletesen ez azt jelenti kétváltozós esetben, hogy az ellipsziseket igyekeztünk körösiteni, és igy a következő pontban a gradiens már határozottabban a minimumhely felé mutat.

A Fletcher-Powell-Davidon módszer esetén kipróbáltuk ezt a gyorsitási lehetőséget ugyancsak az (A) -(E) próbafüggvényekre. A 3. táblázat ezeket az eredményeket mutatja.

Próba fv.	Induló pont	Min. érték	Minimum hely		Költség fv.	Futt.idő /sec/	Iteráció szám
			x_1 x_2	x_3 x_4			
A	(-1,2,1)	0.000	1.000000 1.000000		123	0.250	22
B	(-1.2,1)	0.0000	1.000000 1.000000		58	0.251	11
C	(-3,-1, 0,1)	$0.4135.10^{-8}$	0.006826 -0.000682	0.002970 0.002971	203	0.905	23
D	(1,2 2,2)	$0.6889.10^{-13}$	-0.002509 0.997827	1.000000 1.000001	342	2.487	35
E	(20,20, 20,20)	-2.8654	0.153846 1.500000	-0.384615 0.692307	63	0.283	9

3. táblázat

A 2. és a 3. táblázat megfelelő költségfüggvényeinek összehasonlitásából látható, hogy ezzel a gyorsitási lehetőséggel a Quasi-Newton módszerek valóban gyorsithatók. A fentiekben leirt gyorsitási lehetőségből következik, hogy az eljárás olyan próbafüggvényekre kevésbé hatásos, amelyekre pl. kétváltozós esetben a függvényérték szintvonalak kör, illetve közel kör alakuak.

Végül megjegyezzük, hogy az egyes módszerekre a költségfüggvény nagyjából egyenesen arányos a felhasznált gépidővel, tehát a költségfüggvény a tényleges költségek mértékeként használható.

Jelölések

A, H_i	mátrix
$\underline{x}$	oszlopvektor
T	transzponálás
$(\underline{x}, \underline{y})$	skalársorozat
$\|\|.\|\|$	norma
$\|.\|$	abszolut érték
$f(\underline{x})$	minimalizálandó függvény
$\underline{g}(\underline{x})$	$f(\underline{x})$ deriváltja
α, β, γ	skalárok

Irodalomjegyzék

1. Abaffy J. /1975/ "A Quasi-Newton módszerek egy uj három paraméteres osztálya". Alkalmazott Matematikai Lapok /elfogadva/

2. Abaffy J. /1975/ "Néhány kvadratikusan konvergens feltétel nélküli függvényminimalizáló módszer" Alkalmazott Matematikai Lapok /elfogadva/

3. Abaffy J. /1975/ "A dual-mátrix módszerek egy osztályának leirása" Alkalmazott Matematikai Lapok /leadva/

4. Abaffy J. /1975/ "A Smith módszer és a Fox-Wilkinson módszer ekvivalenciája" Alkalmazott Matematikai Lapok /leadva/

5. Abaffy J. - Sloboda F. /1975/ "General algorithms generating conjugate directions and their applications" Algoritmy vo vypoctovej technike, Slovenska vedeckotechnisa spolocnost 97-108

6. * Brent R.P. /1973/ "Algorithm for Minimisation Without Derivatives" Prentice Hall, Englewood Cliffs. New Jersey.

7. Broyden C.G. /1967/ "Quasi-Newton Methods and their Application to Function Minimisation" Math. of Comp. Vol. 21 368-381

8. Broyden C.G. /1970/ "The convergence of single-rank quasi-Newton methods" Math. of.Comp. Vol. 24 365-382

9. Broyden C.G. /1970/ "The Convergence of a Class of Double-rank Minimisation Algorithms 1. General Considerations" J.Inst.Math.Applics. Vol 6 76-90

10. Broyden C.G. /1970/ "The Convergence of a Class of Double--rank Minimisation Algorithms 2. The new Algorithm" J.Inst.Maths.Applics Vol 6 221-231

11. Broyden C.G., Dennis J.E. Jr. Moré J.J. /1973/ "On the local and superlinear convergence of Quasi-Newton methods" J.Inst.Math.Applics Vol 12 223-245

12. Dixon L.C.W. /1972/ "Quasi-Newton algorithms generate indentical points" Math.Progr. Vol 2 383-387

13. Davidon W.C. /1959/ "Variable Metric Methods for Minimisation" Argonne National Laboratory Report No ANL-5990

14. Evans D.J. /ed/ /1974/ "Software for Numerical Mathematics" Academic Press. London and New York

15. Fletcher R. - Powell M.J.D. /1963/ "A rapidly convergent descent method for minimisation" The Computer J. Vol. 6 163.

16. Fletcher R. - Reeves C.M. /1964/ "Function minimisation by conjugate gradients" The Computer J. Vol 7 149.

17. Fletcher R. /1965/ "Function minimsation without evaulating derivatives-a review" Comp.J. Vol 8 33-41

18. Fox L., Huskey H.D., Wilkinson J.H. /1948/ "Notes on the solution of algebraic linear simultaneous equations" Quart J.Mech.Appl.Math. 1. 149-173

19. ×Gill P.E., Murray W. /ed/ /1974/ "Numerical Methods for constrained Optimization" Academic Press London New York

20. ×Himmelblau D.M. /1972/ Applied Nonlinear Programming

21. Hestenes M.R. - Stiefel E. /1952/ "Methods of Conjugate Gradients for Solving Linear Systems" Journal of Research of the National Bureau of Standards Vol 49 409-436

22. Hooke R. - Jeeves F.A. /1959/ "Direct Search Solution of Numerical and Statistical Problems" Journal of ACM Vol 8 p. 212

23. Huang H.Y. /1970/ "Unified Approach to Quadratically Convergent Algorithms for Function Minimization" JOTA Vol 5 405-423

24. Huang H.Y. /1974/ "Method of Dual Matrices for Function Minimization" JOTA Vol 13 519-537

25. ×Kowalik J. - Osborne M.R. "Methods for unconstained optimization problems" American Elsevier Publishing Company, Inc New-York 1968

26. ×Lootsma /ed/ /1972/ "Numerical methods for non-linear optimization" London: Academic Press

27. × Murray W. /1972/ "Numerical Methods for Unconstrained Optimisation" Academic Press, London and New-York

28. Nelder J.A. - Mead R. /1965/ "A simplex method for function minimization" Compl.J.Vol 7 p. 308-313

29. ×Ortega J.M. - Rheinboldt W.C. /1970/ "Iterative solution of nonlinear equations in several variables" New-York London: Academic Press

30. Overholt K.J. /1965/ "An instability in the Fibonacci and Golden section search methods" B.I.T. Vol 5 284-286

31. Overholdt K.J. /1967/ "Note on Algorithm 2. Fibonacci search, algorithm 7. MINX, and the Golden Section search" Comp.J. Vol 9 414

32. Pearson J.D. /1969/ "Variable metric methods of minimisation" Comp. J. Vol 12 p. 171-178

33. Powell M.J.D. /1962/ "An interative method of finding the minimum of a function of several variables" The Computer J. Vol 5 147.

34. Powell M.J.D. /1964/ "An efficient method of finding the minimum of a function of several variables without calculating derivatives" The Computer J. Vol 7 155-162

35. Powell M.J.D. /1971/ "On the convergence of the variable metric algorithm" J.Inst.Math.Applic. Vol 7 21-36

36. Powell M.J.D. /1964/ "A method for minimizing a sum of squares of non-linear function without calculating derivatives" The Compl.J. Vol 7 303-307

37. Powell M.J.D. /1970/ "A survey of numerical methods for unconstained optimization" SIAM Review Vol 12 79-97

38. Powell M.J.D. /1971/ "Recent advances in unconstained optimization" Math.Prog. 1. 26-55

39. *Rall L.B. /1971/ "Nonlinear Functional Analysis and Applications" Academic Press, London-New-York

40. Rosenbrock H.H. /1960/ "An automatic method for finding the greatest or the least value of a function" The Computer J. Vol 3 175

41. Shah B.V. - Buechler R.J. - Kempthorne O. /1964/ "Some algorithms of minimizing a function of several variables" J.Soc.Indust.Appl.Math. Vol. 12 74-92

42. Smith C.S. /1962/ "The automatic computation of maximum likelihood estimates" N.C.B. Scientific Dept. Report S.C. 846/MR/40

43. Stewart G.W. /1967/ "A modification of Davidon's minimization method to accept difference approximations of derivatives" J. Assoc.Comput. Mach. Vol 14 72-83

44. Zangwill W.J. /1967/ "Minimizing a function without calculating derivatives" The Comp. Jour. Vol. 10 293-296

45. А.Н.Тихонов, В.Я.Арсенин "Методы решения некорректных задач" Изд.Наука, Москва, 1974

A TANULMÁNYOK sorozatban eddig megjelentek:

1/1973 Pásztor Katalin: Módszerek Boole-függvények minimális vagy nem redundáns, $\{\wedge, \vee, \neg\}$ vagy {NOR} vagy {NAND} bázisbeli, zárójeles vagy zárójel nélküli formuláinak előállitására

2/1973 Вашкеви Иштван: Расчленение многосвязных промышленных процессов с помощью вычислительных машин

3/1973 Ádám György: A számitógépipar helyzete 1972 második felében

4/1973 Bányász Csilla: Identification in the Presence of Drift

5/1973* Gyürki J.-Laufer J.-Girnt M.-Somló J.: Optimalizáló adaptiv szerszámgépirányitási 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 Klafszky Emil: Geomatriai programozás és néhány alkalmazása

9/1973 R. Narasimhan: Picture Processing Using 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-számitás egy uj módszeréről

12/1973 Makroprocesszorok, programozási nyelvek. Cikkgyüjtemény az NJSzT és SzTAKI közös kiadásában.
Szerkesztette: Legendi Tamás

13/1973 Jedlovszky Pál: Uj módszer bonyolult rektifikáló oszlopok vegyészmérnöki számitására

14/1973 Bakó András: MTA kutatóintézeteinek bérszámfejtése számitógéppel

15/1973 Ádám György: Kelet-nyugati kapcsolatok a számitógépiparban

16/1973 Fidrich I.-Uzsoky M.: LIDI-72 listakezelő rendszer a Digitális Osztályon, 1972.évi változat

17/1974 Gyürki József: Adaptiv termelésprogramozó rendszer /APS/ termelőmühelyek irányitására

18/1974 Pikler Gyula: MINI-számitógépes interaktiv alkatrészprogramiró 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 telitett mágneses körök számitógépes tervezési 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 József: Numerikus módszerek sparse mátrixokra

27/1974 Somló János: Analitikus optimalizálás

28/1974 Vámos Tibor: Tárgyfelismerési kisérlet nyelvi módszerekkel

29/1974 Móricz Péter: Vegyészmérnöki számitási módszerek fázisegyensulyok és kémiai egyensulyok 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 irányitása különös tekintettel a Fővárosi Gázmüvek 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 Indentification Method for Industrial Robots by Extracting the Main Features of Objects

35/1975 Garádi-Krámli-Ratkó-Ruda: Statisztikai és számitástechnikai módszerek alkalmazása kórházi morbiditás vizsgálatokban

36/1975 Renner Gábor: Elektromágneses tér számitása nagyhőmérsékletü anyagban

37/1975 Edgardo Felipe: Specification problems of a process control display

38/1975 Hajnal Andrásné: Nemlineáris egyenletrendszerek megoldási módszerei

39/1975* A.Abd El-Sattar: Control of induction motor by three phase thyristor connections in the secondary circuit

40/1975 Gerhardt Géza: QDP Grafikus interaktiv 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 M.: Fejezetek a matematikai statiszkikából számitógépes alkalmazásokkal

43/1975 Matavovszky Tibor- dr. Pásztorné Varga Katalin: Programrendszer Boole-függvény együttes egyszerüsitésére vagy minimalizálására

44/1975 Bacsó Nándorné: Pneumatikus áramköri hazardok

45/1975 Varga András: Ellenpárhuzamos félvezetőpárokkal vezérelt aszinkronmotoros hajtások számitási módszerei

46/1976 Galántai Aurél: Egylépéses módszerek lokális hibabecslései

Jelen dolgozat az 5.9.4 számu intézeti témában került kidolgozásra

A x -gal jelölt kivételével a sorozat kötetei megrendelhetők az Intézet könyvtáránál /Budapest, XIII. Victor Hugo u. 18-22/