## (5) Metodă pentru modele standard

Deși la paragrafele precedente s-a discutat repercusiunea cerinței ca $F \in \mathcal{F}^n$ și cum poate fi ea depășită (metodele de prelungire), s-ar putea pune problema de a găsi un algoritm care să nu reclame cerința ca $f_1, f_2$ să fie surjective (v fig 40) Aceasta cu atît mai mult cu cît am demonstrat deja existența și unicitatea soluției pentru exemplul nostru, fără a pretinde ca $f_1, f_2$ să fie surjective (Modelul standard a fost astfel acceptat)

Metoda pe care o vom folosi își are suportul în analiza atentă a teoremei 2, [art 13]:

$\boxed{\text{Teorema 6}}$ (Sandberg)

Pentru ecuația $A\,F(v) + B\,v = C$ (161), unde

1) $f_i$ sunt nedescrescătoare continui $R \xrightarrow{in} R$

2) Matricile A și B satisfac relațiile

$$a_{jj} > \sum_{i \neq j} |a_{ij}| \qquad (181)$$

$$b_{jj} > \sum_{i \neq j} |b_{ij}| \qquad (182)$$

(adică A și B sunt tare dominante pe coloane)

<u>Atunci</u> algoritmul descris mai jos e sigur convergent la soluție.

<u>Algoritmul</u>

Se va proceda în două etape:

1) Se calculează $y^{(n)}$

$$(183)\quad y^{(n)} = -(A - \operatorname{diag} A)\, F(x^{(n-1)}) - (B - \operatorname{diag} B)\, x^{(n-1)} - C$$

2) Se calculează $x^{(n)}$ din ecuația:

$$(184)\quad y^{(n)} = \operatorname{diag} A \cdot F(x^{(n)}) + \operatorname{diag} B \cdot x^{(n)}$$

---

<u>Observații</u>

Este evidentă deosebita libertate lăsată asupra funcțiilor $f_i$. Prețul acestei libertăți (care ne interesează foarte mult dacă vrem să putem considera diodele standard) este apariția condițiilor (181) (182) de dominanță tare pe coloane.

Să vedem dacă putem plăti acest preț în problema exemplu.

$$A = T\begin{pmatrix} 1 & -\alpha_r \\ -\alpha_f & 1 \end{pmatrix} \qquad \begin{matrix} a_{11} > |a_{21}| \\ a_{22} > |a_{12}| \end{matrix} \quad \text{evident.}$$

$$B = G = \begin{pmatrix} G_3 & -G_3 \\ -G_3 & G_1 + G_2 + G_3 \end{pmatrix}$$

$$a_{22} > |a_{12}|$$

<u>dar</u> $\quad a_{11} \not> |a_{12}| \quad (G_3 = |-G_3|)$

Metoda nu poate fi deci aplicată cu certitudine. Ce vom face, în această situație? În cele ce urmează s-a descris o metodă de adaptare a problemei la TG, care urmărește:

— să descrie metoda de adaptare

— să urmărească ce condiție ar trebui îndeplinită pentru ca această adaptare să reușească.

- legat de aceasta, să facă firească introducerea clasei $\mathcal{D}$, în capitolul $\tilde{\text{V}}$,

- să urmărească importanța condiției $\det O \neq 0$ în problemă (legat de aceasta vezi și teoremele de existență și unicitate).

- să familiarizeze pe cititor cu matricile tari dominante pe coloane și proprietățile lor.

Metodă de adaptare

Dacă ecuația $T F(v) + G v = B$ nu a permis folosirea $T G$, ne întrebăm dacă nu cumva s-ar putea face următorul artificiu:

înmulțim ecuația în ambele părți cu o matrice diagonală $D = \begin{pmatrix} d_1 & 0 \\ 0 & d_2 \end{pmatrix}$ (185) cu $\underline{d_1 > 0,\ d_2 > 0}$

Cum $\det O = d_1 d_2 \neq 0$, în acest mod nu influențăm în nici un fel soluțiile ecuației, și obținem:

$$\boxed{DT\,F(v) + DGv = DB} \quad (186)$$

sau

$$A\,F(v) + Bv = C \quad (187) \quad \text{cu} \quad \boxed{\begin{matrix} DT = A \\ DG = B \end{matrix}} \quad (187')$$

Acum, ne putem gîndi la aplicarea teoremei 6, pentru noile matrici $A, B$. Trebuie să verificăm deci că $A$ și $B$ sunt tari dominante pe coloane.

Ca atare ar trebui ca:

1) $A = DT$ tare dominantă pe coloane:

$$\begin{pmatrix} d_1 & 0 \\ 0 & d_2 \end{pmatrix} \cdot \begin{pmatrix} 1 & -\alpha_r \\ -\alpha_t & 1 \end{pmatrix} = A = \begin{pmatrix} d_1 & -d_1\alpha_r \\ -d_2\alpha_t & d_2 \end{pmatrix}$$

și condițiile sînt:

$a_{11} > |a_{21}|$ : $d_1 > |-d_2\alpha_t|$, $\frac{d_1}{d_2} > \alpha_t$

$a_{22} > |a_{12}|$ : $d_2 > |-d_1\alpha_r|$, $\frac{d_2}{d_1} > \alpha_r$ sau:

$$\boxed{\alpha_t < \frac{d_1}{d_2} < \frac{1}{\alpha_r}} \quad (188)$$

Desigur că pot fi aleși astfel de $d_1, d_2$, avînd în vedere faptul că $\alpha_r, \alpha_t \in (0,1)$

Observații

Acum problema se pune astfel: Pot găsi o matrice $D$, care să satisfacă relația (188) și în plus $DG$ să fie tare dom. pe coloane? Acestea sînt cazurile notate cu $\boxed{\mathcal{D}(T)}$ (clasa matricilor $G$, aî se respectă 188, pentru un anumit $T$ și există $D$ aî în același timp $DG$ e tare dominantă pe coloane). Secțiunea de față arată unde pot apare aceste tipuri ($\mathcal{D}(T)$) de matrici $G$ și în ce constă importanța lor (cât de exemplu ele permit aplicarea algoritmului ~~exact~~ al $TG$, în condiții satisfăcătoare pentru funcțiile $t_i$).

2) $B = DG$ este tare dominantă pe coloane.

Observații

Înainte de a trece la calcul să observăm un aspect: anume că dacă $\det G = 0$, atunci niciodată $DG$ nu poate fi tare dominantă pe coloane.

Aceasta dovedeşte $\det G = 0 \Rightarrow \det(DG) =$

$= \det D \cdot \det G = 0$, deci $\forall D > 0$ (matrice diagonală cu elemente pozitive), $DG$ este singulară.

( Putem obţine această concluzie şi astfel (mai lung, dar instructiv din pt. de vedere a modului cum acţionează o înmulţire la dreapta cu o matrice diagonală)

$$\begin{pmatrix} d_1 & & \\ & \ddots & \\ & & d_n \end{pmatrix} G = \begin{pmatrix} d_1 g_{11} & d_1 g_{12} & \dots & d_1 g_{1n} \\ d_2 g_{21} & d_2 g_{22} & \dots & d_2 g_{2n} \\ - & - & - & - \\ d_n g_{n1} & d_n g_{n2} & \dots & d_n g_{nn} \end{pmatrix}$$

$\Rightarrow \det DG = d_1 \cdot d_2 \cdot \ldots \cdot d_n \det G = 0$. )

Acum rămîne să folosim faptul că <u>o matrice tare dominantă pe coloane e nesingulară</u>. Aceasta este T7, BN, dar să arătăm valabilitatea pentru $n = 2$, şi ca o ~~sugestie~~ întărire a demonstraţiei generale. (v [0])

$\det G = G_{11} G_{22} - G_{12} G_{21} \neq 0$,

căci $G_{11} > |G_{21}|$, $G_{22} > |G_{12}|$ $\quad$ ş.a.

---

Revenind la exemplul nostru:

$$\begin{pmatrix} d_1 & 0 \\ 0 & d_2 \end{pmatrix} \cdot \begin{pmatrix} G_3 & -G_3 \\ -G_3 & G_1+G_2+G_3 \end{pmatrix} = \begin{pmatrix} d_1 G_3 & -d_1 G_3 \\ -d_2 G_3 & d_2 (G_1+G_2+G_3) \end{pmatrix}$$

Deci trebuie ca:

$$d_1 G_3 > |d_2 G_3| \Longrightarrow d_1 > d_2 \qquad (185)$$

$$d_2 (G_1+G_2+G_3) > d_1 G_3 \Rightarrow \frac{d_1}{d_2} < \frac{G_1+G_2+G_3}{G_3} \qquad (190)$$

Sau condensat

$$\boxed{1 < \frac{d_1}{d_2} < 1 + \frac{G_1+G_2}{G_3}} \qquad (191)$$

Și această relație permite alegerea unui $d_1$ și $d_2$, mai mult, se pot alege $d_1$ și $d_2$ astfel încît să fie satisfăcută și (188). De exemplu:

$$\boxed{1 < \frac{d_1}{d_2} < \min\left\{\frac{1}{\alpha_T}, 1+\frac{G_1+G_2}{G_3}\right\}} \quad (192)$$

satisface ambele relații (188), (191), deci $DT$, $DG$ sînt tari dominante pe coloane. ~~Sa~~ (deci $(T, G) \in \mathcal{D}$, sau $G \in \mathcal{D}(T)$). Se pot alege astfel de $d_1$, $d_2$, căci $\alpha_T < 1$, $1+\frac{G_2+G_1}{G_3} > 1$ //.

$\boxed{\text{Concluzia 4}}$

<u>Dacă</u> $G_1, G_2, G_3 > 0$, $\alpha_T, \alpha_t \in [0,1]$ și $k_1, k_2$ sunt ~~~~ mediu crescătoare (pt. t. de existență și unicitate pentru forări cf. pag 135), <u>există</u> un algoritm de tip „Sandberg adaptat" (v. mai jos) care converge riguros la soluție.

<u>Algoritm</u>

- se alege un $d_1, d_2$ care să verifice (192) (pentru decizii s-ar putea lua $d_2 = 1$, $d_1 = \frac{1+b}{2}$, unde $b$ este $\min\left\{\frac{1}{\alpha_T}, 1+\frac{G_1+G_2}{G_3}\right\}$.)
- se înmulțește ecuația inițială (92) cu $D$ astfel ales și se găsește (186).
- acestei ecuații i se aplică algoritmul descris clar în relațiile (183), (184).

Observație

Ecuația (184) care intervine este bine definită (are o soluție unică. Fiecare „componentă" a ecuației are forma: $y_k^{(n)} = a_{kk} f(x_k^{(n)}) + b_{kk} x_k^{(n)}$, care este o bijecție, având în vedere proprietatea de monotonie strictă a lui f și că $a_{kk}, b_{kk} > 0$ (v (181), (182)).

## 6 Calculul marginilor soluțiilor și rolul său

Există metode prin care reușim să estimăm [Var. 1] locația soluțiilor, pentru un B fixat, fără a rezolva ecuațiile. Aceste metode sunt cu atât mai valoroase

- cu cât sînt mai simple de aplicat (eventual manual)
- cu cât se aplică în situații mai largi
- atunci când se găsește „zona" U în care se află soluțiile atunci când B variază într-o zonă H (Var 2)

Mai întîi să evidențiem:

a) Rolul calculului marginilor soluțiilor

Pentru ecuația simplă: $F(U) + T^{-1} G U = T^{-1} B$, toți algoritmii pe care i-am propus anterior, pornesc de la un $x_0$ oarecare. Prima utilitate evidentă este deci:

U1 : Estimarea lui $x_0$ și prin aceasta pornirea „în forță" a algoritmului de calcul (ec. de timp și bani)

Mai mult, dacă algoritmul folosit nu este convergent în întreg spațiul propus, este posibil ca, plasarea problemei în „zona" unde se află soluțiile:

U2 : Să facă unii algoritmi necesari convergenți

La aceasta să adăugăm și un lot de aspecte legate de probleme de calitate, care au fost deja menționate la locul respectiv

U3 : Implicațiile în problema „intrare mărginită - - ieșire mărginită" și prin aceasta, în coroborare cu teoremele lui Palais; (v pag 148)

U4 : implicații în problema existenței și unicității soluțiilor (v pag 148)

U5 : ridicarea în rang (practic) al problemei (pag 149)

Dar mai ales trebuie insistat asupra unui aspect, menționat anterior (pag 148 β) - și pag 137. - fig 36. și anume:

U6 : Posibilitatea analizei cazului când funcțiile au domeniu ~~...~~ mărginit, prin <u>procedeul de prelungire</u> (v fig 36) urmat de <u>verificarea de coincidență</u>, care constă tocmai în aceea că se obțin marginile între care se află soluțiile ecuațiilor „prelungite" și se verifică dacă

Funcțiile inițiale au un domeniu care cuprinde aceste margini. In acest context, și dacă avem grijă ca procedeul de prelungire să nu altereze aspectul calitativ al funcțiilor (deci de exemplu să dispunem de o teoremă de existență și unicitate pentru funcțiile prelungite) se poate arăta (v. [0]) și este destul oarecum intuitivă

Propoziția de prelungire (T.7) Dacă pentru o ecuație, conform unui procedeu oarecare stabilim că soluția se află într-o anumită zonă H și e unică, funcțiile respective pot fi modificate, cu condiția ca ele să rămână la fel în zona H, iar prin modificarea să nu se cauzeze o conformație calitativ nouă (care să admită apariția altor soluții)

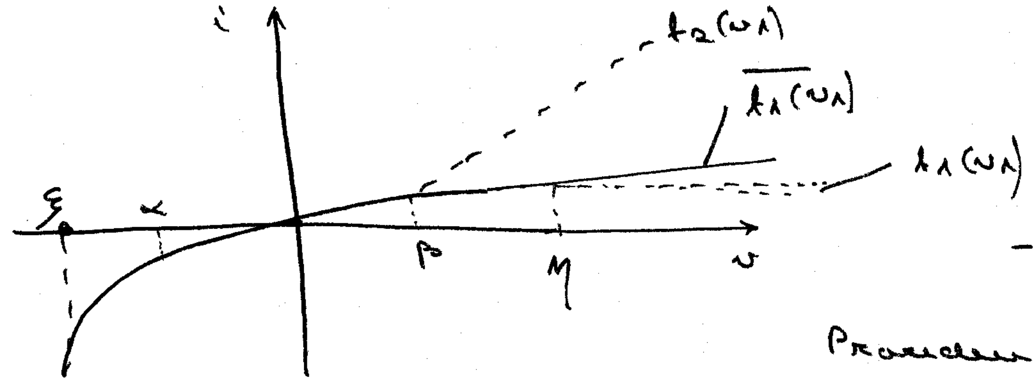
U7 Procedând ca mai sus, după ce aflăm marginile soluțiilor ecuațiilor originale, putem apela la o pre modificare de tipul de mai sus cu efecte pozitive:

— să facă posibil un algoritm care nu era aplicabil ecuațiilor originale

— să mărească viteza de convergență a unui algoritm.

Un exemplu în acest sens ni-l oferă conclusia 1 (pag 164). Se vede că este necesară ca panta funcțiilor $f_i$ să fie mărginită inferior de $\varepsilon$. (fig 40) pag 162

Ori, pentru cazul diodelor standard, $f_1(v_1)$ din




— fig 42 —

Procedeu de modificare („prelungire)

fig 42, deși panta e pozitivă, ea se apropie oricât de 0. De aceea condiția 1 nu poate fi aplicată. Dar, dacă în prealabil determinăm, pentru ecuația originală că are o soluție unică între $\xi$ și $\eta$, atunci putem face „prelungirea acceptabilă" $\overline{f_1(v_1)}$ (desigur nu vom adăuga nici-o soluție nouă, $\overline{f_1}(v_1)$ și $f_1(v_1)$ satisfac ambele cerințele de monotonie suficiente pentru a deduce existența și unicitatea în problema noastră). Funcția $\overline{f_1}(v_1)$ satisface însă cerința de pantă a teoremei.

Mai mult, dacă se obține mărginirea $(\alpha, \beta)$, putem chiar modifica funcția și între $\beta$ și $\eta$ ($\overline{f_2(v_1)}$) căci în acest mod ~~f~~ panta minimă crește și vom vedea (7) că o dată cu ea crește și viteza de convergență.

b) Teorema de calcul al mărginilor (Varianta 1 – fixă)

$\boxed{\text{Teorema 8}}$

**Dacă** ecuația (193) $F(x) + Ax = B$ satisface condițiile:

(194) 1) $f_i \in \mathcal{F}$

(195) 2) $a_{ii} \geq \sum_{\substack{j=1 \\ j\neq i}}^{n} |a_{ij}|$ (slabă dom. pe linii)

**atunci** soluția lui (147) se găsește în domeniul $\underline{R}$ obținut prin algoritmul de mai jos

**Algoritm**

1) – se rezolvă ecuația $F(x) = B$ (196)

sau $\begin{cases} f_1(x_1) = b_1 \\ f_2(x_2) = b_2 \\ \vdots \\ f_n(x_n) = b_n \end{cases}$ (197) $\xrightarrow{\text{soluțiile}}$ $\begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix}$ (198)

2) – se calculează: $\alpha = \max\{|\alpha_i|\}$ (199)

și $B' = \left[\sum_{j\neq 1}^{n} |a_{ij}|, \sum_{j\neq 2}^{n} |a_{ij}|, \ldots \sum_{\substack{j\neq n \\ j=1}}^{n-1} |a_{ij}|\right]^t$ (200)

3) Se rezolvă ecuațiile:

(201) $F(x) + \mathrm{diag}\, A \cdot x = B - \alpha B'$

(202) $F(x) + \mathrm{diag}\, A \cdot x = B + \alpha B'$, notându-se soluțiile respectiv cu

(203) $\eta = [\eta_1, \ldots \eta_n]^t$ pentru (201)

(204) $\xi = [\xi_1, \ldots \xi_n]^t$ pentru (202)

4) Atunci domeniul soluțiilor este

$R = i_1 \times i_2 \times \ldots \times i_n$ cu $i_k = [\eta_k, \xi_k]$ $k = 1 \ldots n$

(205)

## Exemplu de aplicare

Pentru ecuația următoare:

$F(v) + T^{-1} G v = B_1 = T^{-1} B$ sau pe larg:

$$F(v) + \frac{1}{1-\alpha_r\alpha_t}\begin{pmatrix}1 & \alpha_r\\ \alpha_t & 1\end{pmatrix}\begin{pmatrix} G_3 & -G_3 \\ -G_3 & G_1+G_2+G_3\end{pmatrix} = \frac{1}{1-\alpha_r\alpha_t}\begin{pmatrix}1&\alpha_r\\ \alpha_t & 1\end{pmatrix}\begin{pmatrix}-G_3\\ G_2+G_3\end{pmatrix}E_c \qquad (206)$$

observăm în primul rînd că se cer aceleași condiții ca la 72 pag 159, în afară de cea de pantă mărginită inferior uniform pozitiv (145).

Presupunem așadar că $f_1, f_2 \in \mathcal{F}$ și că

(207) $\boxed{G_3 \geq \frac{\alpha_r}{2(1-\alpha_r)}(G_1+G_2)}$ (v. Concluzia 1, pag 164) și sîntem acum siguri de îndeplinirea condițiilor ($\boxed{\text{Consecința 5}}$)

Mersul calculului:

1) Se rezolvă ecuația

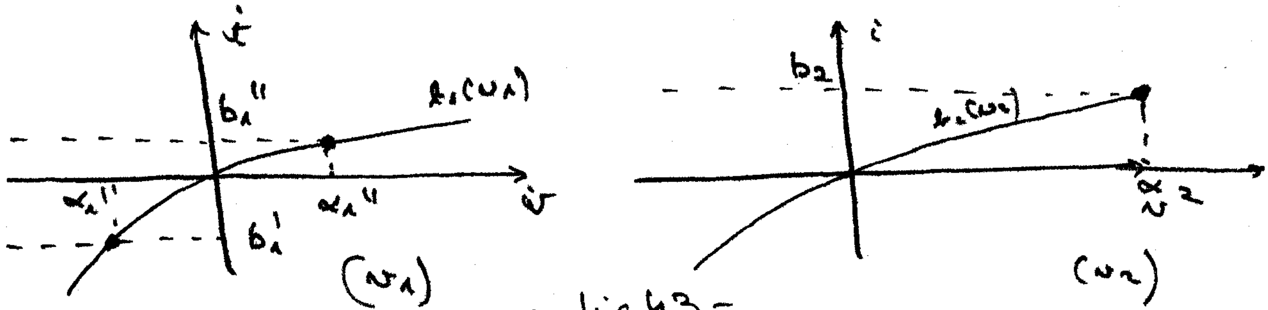
$$\begin{cases} f_1(v_1) = b_1 = \dfrac{E_c}{1-\alpha_r\alpha_t}\left[-G_3+\alpha_r(G_2+G_3)\right] \\ f_2(v_2) = b_2 = \dfrac{E_c}{1-\alpha_r\alpha_t}\left[-\alpha_t G_3 + G_2+G_3\right]\end{cases}$$

cu $\frac{1}{1-\alpha_r\alpha_t} = K$ aceasta se mai scrie:

$$\begin{aligned} f_1(v_1) &= KE_c\left(G_3(\alpha_r-1)+\alpha_r G_2\right) = b_1 \\ f_2(v_2) &= KE_c\left[G_3(1-\alpha_t)+G_2\right] = b_2 \end{aligned} \qquad (208)$$

Se vede că $b_1 < b_2$ ($b_2$ e totdeauna pozitiv, $b_1$ poate fi și negativ dacă $G_3(1-\alpha_r) > \alpha_r G_2$, adică $G_3 > \frac{\alpha_r}{1-\alpha_r} G_2$).

Grafic deci situația arată:




— fig 43 —

şi se vede că în general $b_1' \geq 0$, $b_2 > 0$ duce la soluţiile prezentate. Dacă $f_1, f_2$ nu diferă prea mult, $\alpha_2$ va fi cel mai mare în modul. Oricum ar fi,

2) Se ia $\alpha = \max\{|\alpha_1|, |\alpha_2|\}$ (209)

2) Se calculează

$$B' = \begin{bmatrix}|a_{12}|\\|a_{21}|\end{bmatrix} = \begin{bmatrix}|G_3(\kappa_0 - 1) + \kappa_0(G_1+G_2)| \\ G_3(1-\kappa_t)\end{bmatrix} \qquad (209')$$

(cu atenţia de rigoare asupra modulului lui $|a_{12}|$)

3) se rezolvă ecuaţiile:

$$\begin{pmatrix} f_1(v_1) \\ f_2(v_2)\end{pmatrix} + \begin{pmatrix} a_{11} & 0 \\ 0 & a_{22}\end{pmatrix}\begin{pmatrix} v_1 \\ v_2\end{pmatrix} = B - \alpha B'$$

$$f_1(v_1) + a_{11} v_1 = B_1 - \alpha B_1' \longrightarrow v_1 = \xi_1 \qquad (210)$$

$$f_2(v_2) + a_{22} v_2 = B_2 - \alpha B_2' \longrightarrow v_2 = \xi_2 \qquad (211)$$

$$\begin{pmatrix} f_1(v_1) \\ f_2(v_2)\end{pmatrix} + \begin{pmatrix} a_{11} & 0 \\ 0 & a_{22}\end{pmatrix}\begin{pmatrix} v_1 \\ v_2\end{pmatrix} = B + \alpha B'$$

$$f_1(v_1) + a_{11} v_1 = B_1 + \alpha B_1' \longrightarrow v_1 = \eta_1 \qquad (212)$$

$$f_2(v_2) + a_{22} v_2 = B_2 + \alpha B_2' \longrightarrow v_2 = \eta_2 \qquad (213)$$

4) Cu valorile găsite mai sus, se calculează

$$\boxed{R = i_1 \times i_2 = (\xi_1, \eta_1) \times (\xi_2, \eta_2)} \qquad (214)$$

<u>Comentarii</u>

1) Din nou am putea justifica la [illegible]

pas buna definire a procedeului. (ec. 208, 210, 211, 212, 213)

Nu mai fac, ~~pentru că~~, (de altfel se rezultă și din concluzia teoremei)

2) ecuațiile implicate la pașii 1 și 4 sînt foarte simple. {

Ecuațiile $(208)_1$ au, pentru partea exponențială o soluție logaritmică, iar pentru partea liniară (prelungirea eventuală) una liniară.

Ecuațiile (210) → (213) pot fi rezolvate prin metode iterative.

Merită însă să observăm un fapt: dacă admitem o ușoară lărgire a marginilor obținute, putem obține un

$$I' = (\xi_1', \eta_1') \times (\xi_2', \eta_2')$$

cu $\xi'_{1(2)} < (\text{puțin})\, \xi_{1(2)}$ $\qquad \eta'_{1(2)} > (\text{puțin})\, \eta_{1(2)}$

<u>foarte ușor</u>, manual, prin aproximații ~~[illegible]~~, de genul:

Fie $h(v_1) = t_1(v_1) + \sigma_3(1-\alpha r)v_1 - \bar{b}_1$

Se încearcă un $v_1^0$, dacă $h(v_1^0) < 0$ se caută un $v_1^1$ mai mare, pînă cînd $h(v_1^1) \geq 0$. ~~Pot alege~~ Acum în intervalul $(v_1^0, v_1^1)$ aplic metoda înjumătățirii, de cîteva ori, urmărind semnul pînă mă apropii suficient de mult de soluția exactă (detaliile sînt evidente).

Astfel vom căuta, în ecuațiile (210) (211) o margine „rezonabil de apropiată inferior de soluții, iar în (212) (213) margini apropiate superior.

3) După ce am găsit marginile, aplicăm metoda de prelungire sugerată la pag 190 (discuția s-a făcut acolo). Ca urmare

$\boxed{\text{Consecința 6}}$

În condițiile consecinței 5 (pag 192) putem ca după ce am găsit pe $(q_1, \gamma_1)$, $(q_2, \gamma_2)$ să modificăm funcțiile $f_1, f_2$ ca la pag 190 (dincolo de $\gamma_1$ și $\gamma_2$ funcția se duce cu pantă $\geq \varepsilon$ - v. fig 42, modificarea $\overline{f_2(u_1)}$)

Astfel, putem aplica automat după aceea varianta nr. 1, algoritmul de la consecința 1 (pag 164), ajungând riguros la soluția corectă.

ⓒ Teorema de calcul al marg. - Var-2 (variabilă)

În condițiile teoremei 1, dacă ni se dau limitele în care „joacă" B:

$$(215) \begin{cases} \alpha_1 \leq b_1 \leq \beta_1 \\ \alpha_2 \leq b_2 \leq \beta_2 \\ \vdots \\ \alpha_n \leq b_n \leq \beta_n \end{cases}$$

sau pe scurt:

$$(215') \quad \alpha \leq B \leq \beta$$

(vectori coloană)

se stabilesc limitele între care se află soluția printr-un procedeu cu totul analog celui precedent.

→ Aceasta este $\boxed{\text{Teorema 9}}$, iar algoritmul este dat explicit în continuare:

Algoritm

- se rezolvă ecuațiile

$$F(\alpha) = \alpha \qquad (216)$$

$$F(x) = \beta \qquad (216')$$

a căror soluții se notează

$$\mu = (\mu_1 \ldots \mu_n)^t \qquad (217)$$

$$\nu = (\nu_1, \ldots \nu_n)^t \qquad (217')$$

- se calculează

$$B' = \left[ \sum_{j \neq 1} |a_{1j}|, \ldots \sum_{j=1}^{n-1} |a_{nj}| \right]^t \qquad (218)$$

și

$$\lambda = \max\{|\mu_1|, |\mu_2|, \ldots |\mu_n|, |\nu_1| \ldots |\nu_n|\} \qquad (219)$$

- se rezolvă ecuațiile:

$$F(x) + \operatorname{diag} A \cdot x = \alpha - \lambda B' \qquad (220)$$

$$F(x) + \operatorname{diag} A \cdot x = \beta + \lambda B' \qquad (221)$$

notându-le soluțiile cu

$$\eta = [\eta_1, \ldots \eta_n] \qquad (222)$$

$$\xi = [\xi_1, \ldots \xi_n] \qquad (223)$$

- Atunci:

$$\begin{cases} \eta_1 \leq x_1 \leq \xi_1 \\ \text{------} \\ \eta_n \leq x_n \leq \xi_n \end{cases} \quad \text{sau} \qquad (224)$$

$$x \in R = i_1 \times i_2 \ldots \times i_n \quad \text{cu} \quad i_k = [\eta_k, \xi_k] \quad k = 1 \ldots n$$

Observații

În cadrul demonstrației [vezi [7]] se arată și faptul că ecuația (225) $F(x) + \operatorname{diag} A x = B$, oferă și ea niște soluții, să le notăm cu $x^*$, cu proprietatea că se află în intervalele $[\eta_k, \xi_k]$ găsite mai sus sau pentru cazul teoremei 8, între soluțiile ecuațiilor (201), (202). Mai mult se arată că dacă matricea A devine diagonală (adică $\sum_{j\neq i} a_{ij} \to 0$), atunci soluțiile ecuațiilor (201)(202) se strâng la $x^*$, și deci o dată cu ele și soluția reală a ecuațiilor. Toate acestea sunt motive suficiente pentru:

$\boxed{\text{Consecința 7}}$

Rezolvând ecuația (225), vom găsi o soluție $x^*$, care este o bună aproximație inițială, (de start) a algoritmului.

2) Se mai arată că $\alpha$ obținut la 199 este o majorare a marginilor, deci că

$$-\alpha \leq \eta_i \leq \xi_i \leq \alpha \qquad \forall i = 1, \dots n \quad \text{deci}$$

$\boxed{\text{Consecința 8}}$ :

Putem obține o primă estimare (corectă) a zonei în care se află soluțiile, rezolvând ecuația $F(x) = B$ (197) =)(198) $[\alpha_1, \dots \alpha_n]^T$ =) $\alpha = \max\{|\alpha_i|\}$ și atunci $H = i_1 \times i_2 \times \dots \times i_n$ cu $i_k = [-\alpha, \alpha]$.

Exemple:

(Cons 7) Se rezolvă ecuația

$$(226) \quad \begin{cases} k_1(v_1) + a_{11} v_1 = b_1 \\ k_2(v_2) + a_{22} v_2 = b_2 \end{cases}$$ , obținând un

punct $\begin{pmatrix} v_1^* \\ v_2^* \end{pmatrix}^{(227)}$, valabil ca primă aproximare a lui $\begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$ (~~v~~ fig 43)

(Cons. 8) Se rezolvă ecuația:

$$(228)\begin{cases} t_1(v_1) = b_1 \\ t_2(v_2) = b_2 \end{cases} \Rightarrow (229)\begin{cases} v_1 = \alpha_1 \\ v_2 = \alpha_2 \end{cases}$$

, ceea ce este foarte simplu (v fig 43), descriptia de la pag 194)

Avem o primă „majorare" a zonei în care se găsesc soluțiile : $i = [-\alpha, \alpha] \times [-\alpha, \alpha]$. (230)

## © Teoreme general de calcul al marginilor

Problema pusă la acest punct este analoagă celei de la par 4: nu s-ar putea obține metode de calcul al marginilor, paralele teoremelor de ex. și unicitate, care să nu ceară îndeplinirea suplimentară a altor condiții? (analoage Teoremei 5)

Răspunsul este oferit de următoarea teoremă :

$\boxed{\text{Teorema 10}}$ **Dacă**: ecuația $F(x) + Ax = B$ are

1) $F \in \hat{\mathcal{F}}^m$ (231)

2) $A \in P_0$ (232)

atunci fiind dați $\alpha_i \leq \beta_i$ a.î.:

$$\alpha_i \leq b_i \leq \beta_i \quad \forall i = 1, \dots m \quad (233)$$

atunci există numerele $\gamma_i, \delta_i$ $i = 1, \dots m$, a.î.

$$\boxed{\gamma_i \leq x_i \leq \delta_i} \quad \forall i = 1 \dots m \quad (234)$$

și putem să construim efectiv aceste margini

Procedeul va fi ilustrat în cele ce urmează pentru $n = 2$ (probl. exemplu)

Cazul fix se obține din cel variabil, făcând $\alpha_i = \beta_i = b_i$ și nu va mai fi discutat aparte

Concluzia 9 Dacă pentru ecuația exemplu (92)

$$\left( F(v) + A v = \bar{B} \quad \text{cu } A = \tau^{-1} G,\ \bar{B} = \tau^{-1} B \right)$$

(235) — $F \in \tilde{F}^{2n}$

— $A \in P_0$ (care am arătat aceasta pentru $\left\langle \begin{array}{l} G_1, G_2, G_3 > 0 \\ x_r < \ell \in (0,1) \end{array} \right.$)

atunci fiind dați $\alpha_i, \beta_i$ a.î.

$$(236)\ \begin{cases} \alpha_1 < b_1 < \beta_1 \\ \alpha_2 < b_2 < \beta_2 \end{cases}$$

putem găsi $v_1, v_2, \sigma_1, \sigma_2$ a.î.:

$$(237)\ \begin{cases} v_1 \le v_1 \le \sigma_1 \\ v_2 \le v_2 \le \sigma_2 \end{cases}$$ prin procedeul redat mai jos:

Demonstrație - construcție

Adusă în forma:

$$(238)\ F(v) + A v = B \qquad \left( A = \tau^{-1} G,\ B = \tau^{-1} E_c \begin{pmatrix} G_3 \\ -G_2 + G_3 \end{pmatrix} \right)$$

am arătat că $A \in P_0$, înseamnă că

$$(239)\ \det \left\{ \begin{pmatrix} d_1 & 0 \\ 0 & d_2 \end{pmatrix} + A \right\} \neq 0 \quad \text{pentru } d_1, d_2 > 0.$$

condiție îndeplinită pentru exemplul nostru

Să mai observăm că minorii (matrici) de ordin 1 principali ai lui $A$, $[a_{11}]$, $[a_{22}]$ sunt tot în $P_0$, care evident că $a_{11} + d_1 \neq 0$

$a_{22} + d_2 \neq 0$ pt $d_1, d_2 > 0$, căci $a_{11}, a_{22} \geq 0$.

---

T11

O altă observație, care prezintă un interes de sine stătător, este că matricele din clasa $P_0$, mai au și următoarea proprietate: $\forall x$, există un indice $k$ a.î. $x_k$ are același semn cu $(Ax)_k$. (v cap IV B). Prelucrăm de o carte pentru a demonstra această propoziție pentru $n=2$. Acum vrem să arătăm că

$\forall\, v_1, v_2$:

$$(240)\ \begin{cases} v_1(a_{11}v_1 + a_{12}v_2) \geq 0 \\ \quad\text{sau} \\ v_2(a_{21}v_1 + a_{22}v_2) \geq 0 \end{cases}$$ (în sensul că măcar una din relații e îndeplinită)

Demonstrație

Să presupunem că (240) nu este adevărată, deci că

$$(241)\ \begin{matrix} v_1(a_{11}v_1 + a_{12}v_2) < 0 \\ v_2(a_{21}v_1 + a_{22}v_2) < 0 \end{matrix}$$ pentru un $(v_1, v_2)$ (desigur $v_1, v_2 \neq 0$)

Atunci, pentru acest $(v_1, v_2)^T$ vom avea:

$$\left\{\begin{pmatrix} d_1 & 0 \\ 0 & d_2 \end{pmatrix} + \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}\right\}\begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} d_1 v_1 + [a_{11}v_1 + a_{12}v_2] \\ d_2 v_2 + [a_{21}v_1 + a_{22}v_2] \end{bmatrix}$$

și este evident că pentru:

$$(242)\ d_1 = -\frac{a_{11}v_1 + a_{12}v_2}{v_1} \quad (242')\ d_2 = \frac{a_{21}v_1 + a_{22}v_2}{v_2}$$ obținem:

$$\left\{\begin{pmatrix} d_1 & 0 \\ 0 & d_2 \end{pmatrix} + \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}\right\}\begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad (243)$$

Dar observînd că relațiile (241) impun (în 242) că $d_1, d_2 > 0$, relația (243) contravine cerinței (239), pe care am demonstrat-o deja.

Deci (241) nu e posibilă $\Longrightarrow$ (240) q.e.d.

---

Acum să trecem la procedeul de calcul al marginilor: Relația de bază este (240). Vom presupune pe rînd că se realizează una din cele două variante (știm că măcar una se realizează)

(1) $v_1(a_{11}v_1 + a_{12}v_2) \geq 0$ (244)

Atunci

$v_1 t_1(v_1) + v_1(a_{11}v_1 + a_{12}v_2) = v_1 b_1$ duce la (245)

$$\boxed{v_1 t_1(v_1) \leq v_1 b_1} \quad (246)$$

Această relaţie este esenţială în procedura ce urmează. (care trebuie urmărită atent pentru că pe ea se va sprijini fiecare pas al calculului)

$v_1(t_1(v_1) - b_1) \geq 0$ (247), cu două posibilităţi:

$$\begin{cases} v_1 \geq 0 \\ t_1(v_1) \leq b_1 \end{cases} \ (248') \qquad \text{sau} \qquad \begin{cases} v_1 \leq 0 \\ t_1(v_1) \geq b_1 \end{cases} \ (248'')$$

Vom defini (şi calcula)

$$\boxed{\gamma_1^1 = \min\{t_1^{-1}(\alpha_1), 0\} \qquad \delta_1^1 = \max\{t_1^{-1}(\beta_1), 0\}} \quad (249)$$

şi pretindem că $\boxed{v_1 \in [\gamma_1^1, \delta_1^1]}$ (250)

Într-adevăr, se observă de la început că

(251) $\gamma_1^1 \leq 0$
(252) $\delta_1^1 \geq 0$, aşa că, în 247, cazul $v_1 = 0$ este satisfăcând banal relaţia (250).

Dacă $v_1 \neq 0$ în 247, avem posibilităţile:

1) (248') $v_1 \geq 0$, atunci evident $v_1 \geq \gamma_1^1$ (v 251), iar din (248') $\Rightarrow t_1(v_1) \leq b_1 \overset{(236)}{\leq} \beta_1 \Rightarrow v_1 \leq t^{-1}(\beta_1) \leq \delta_1^1$, unde am folosit monotonia lui $t$.

2) Dacă $v_1 \leq 0$, atunci $v_1 \leq \delta_1^1$ (252), iar din (248'')
$t_1(v_1) \geq b_1 \overset{236}{\geq} \alpha_1 \Rightarrow v_1 \geq t^{-1}(\alpha_1) \geq \gamma_1^1$.

În concluzie, în acest caz:

(250) $v_1 \in [\gamma_1^1, \delta_1^1]$, calculați ca mai sus.

În continuare vom încerca să găsim o margine și pentru $v_2$. Pentru aceasta vom pleca de la ecuația defalcată: $f_2(v_2) + a_{21} v_1 + a_{22} v_2 = b_2$ (253)

pe care o vom scrie:

$$f_2(v_2) + a_{22} v_2 = b_2 - a_{21} v_1 \qquad (254)$$

Bazați pe această formă, vom găsi margini pentru $v_2$, deoarece ecuația (254) e de aceeași formă cu (238) ($F_2(x) + A_2 \cdot x = B_2$), cu un ordin mai mic, în care rolul $A_2$, acum $[a_{22}]$ este sigur tot în $P_0$. (vezi începutul demonstrației, pentru cazul particular, și cap IV pentru cel general - ideea e că, tăind o linie și o coloană de același indice, noul determinant va avea toți minorii principali printre cei vechi - ai lui $A$)

Să notăm (255) $b_2 - a_{21} v_1 = \tilde{b}_2$ și să observăm că, dacă

$$\alpha_2 \le b_2 \le \beta_2 \qquad (237)$$

$$\gamma_1^1 \le v_1 \le \delta_1^1 \qquad (250)$$, atunci

vom găsi margini și pentru $\tilde{b}_2$, să le spunem $\tilde{\alpha}_2, \tilde{\beta}_2$, cu cazurile $a_{21} > 0$, $a_{21} < 0$ analoage. La noi $a_{21} = G_3(\kappa_t - 1) < 0$

așa că scriem relația pentru acest caz, obținută prin înmulțirea lui (250) cu $-a_{21} > 0$ și adunarea la 237:

$$\underbrace{\alpha_2 - a_{21}\gamma_1^1}_{= \tilde{\alpha}_2} \le b_2 - a_{21} v_1 \le \underbrace{\beta_2 - a_{21}\delta_1^1}_{= \tilde{\beta}_2} \qquad (256)$$

Acum revenim exact în cazul primei ecuații. Procedăm la fel, o înmulțim cu $v_2$, rezultă:

$$v_2 t_2(v_2) + v_2 a_{22} v_2 = v_2 \tilde{b}_2$$

și cum $v_2 a_{22} v_2 \geq 0$, deducem:

$$v_2 t_2(v_2) \leq v_2 \tilde{b}_2 \qquad (257)$$

Nu mai trebuie să repetăm construcția. Urmăm exact în punctul de la (246) cu

$$\left| \begin{array}{l} v_1 \longrightarrow v_2 \\ t_1 \longrightarrow t_2 \end{array} \right| \begin{array}{l} \alpha_1 \to \tilde{\alpha}_2 \mid \beta_1 \to \tilde{\beta}_2 \\ b_1 \longrightarrow \tilde{b}_2 \end{array}$$

, deci se construiește:

$$\gamma_2^1 = \min\{t_2^{-1}(\tilde{\alpha}_2, 0)\} \qquad \delta_2^1 = \max\{t_2^{-1}(\tilde{\beta}_2, 0)\} \quad (258)$$

și se arată că $v_2 \in [\gamma_2^1, \delta_2^1]$ (259)

Astfel am găsit (250 și 259) margini pentru $v_1$ și $v_2$, în cazul îndeplinirii condiției (244). Mai avem de tratat cazul analog:

2 $v_2(a_{21} v_1 + a_{22} v_2) \geq 0$ (260)

Procedând absolut analog, găsim un nou grup de margini, întâi pentru $v_2$ și apoi pentru $v_1$, adică

$$\left. \begin{array}{l} v_2 \in [\gamma_2^2, \delta_2^2] \\ v_1 \in [\gamma_1^2, \delta_2^1] \end{array} \right) (261)$$

De-abia acum putem spune, cu certitudine că concluzia 9 este valabilă, ajungând în final:

$$\boxed{\begin{array}{ll} \gamma_1 \text{ final} = \min[\gamma_1^1, \gamma_1^2] & \gamma_2 \text{ final} = \min[\gamma_1^2, \gamma_2^2] \\ \delta_1 \text{ final} = \max[\delta_1^1, \delta_1^2] & \delta_2 \text{ final} = \max[\delta_2^1, \delta_2^2] \end{array}} \quad (262)$$

Şi cci $\left| \begin{array}{l} v_1 \in [\gamma_1 t, \delta_1 t] \\ v_2 \in [\gamma_2 t, \delta_2 t] \end{array} \right.$ $\forall$ a.î. $\begin{array}{l} b_1 \in [\alpha_1, \beta_1] \\ b_2 \in [\alpha_2, \beta_2] \end{array}$ (263)

q.e.d.

Rezumatul algoritmului de calcul.

1) Se efectuează mai întîi operaţia (1): „ecuaţia 1 urmată de ecuaţia 2" în ordinea: (249) (255) (256) (258).

2) Se efectuează apoi (2): „ecuaţia 2 urmată de 1", similară primeia (în relaţiile de mai sus se face:

| | | |
|---|---|---|
| $t_1 \to t_2$ | $\alpha_2 \to \alpha_1$ | $t_2 \to t_1$ |
| $\alpha_1 \to \alpha_2$ | $\beta_2 \to \beta_1$ | $\tilde{\alpha}_2 \to \tilde{\alpha}_1$ |
| $\beta_1 \to \beta_2$ | $a_{21} \to a_{12}$ | $\tilde{\beta}_2 \to \tilde{\beta}_1$ |
| (249) | $b_2 \to b_1$ | (în 258) |
| | $\tilde{\alpha}_2 \to \tilde{\alpha}_1$ | |
| | $\tilde{\beta}_2 \to \tilde{\beta}_1$ | |
| | (255, 256) | |

(264)

3) Se face operaţia (262) şi apoi se scrie relaţia finală (263)

Obs: Operaţiile angrenate sunt foarte simple: comparaţii şi calculul lui $t_1^{-1}$, $t_2^{-1}$ în diferite puncte.

Consecinţe (vezi şi discuţia de la pag 190)

Sunt date mai jos exemple de cum poate colabora această ultimă teoremă cu cele de calcul (Conse 1, 2, 3).

În mare principiul este acelaşi:

– se aplică, dacă este posibil (şi am văzut că pentru exemplu nostru nu mai trebuie decît ca $a_1, a_2, a_3 \geq 0$, $\alpha, \alpha_1 \in (0,1)$ şi $t_1, t_2 \in \hat{\sigma}$) algoritmul precedent şi se află marginile (263).

~~Apoi se mo~~

– Apoi se modifică funcţia admisibilă (v 77 pag 189) astfel încât să poată fi aplicați algoritmii de calcul

= În final se face verificarea de coincidență, căci suntem acum siguri că soluția este cea dorită. De exemplu, consecința 6, poate fi urmare a aplicării alg. precedent.

Dar și mai importantă este utilizarea calculului marginilor în probleme de următoarea factură:

— $f_1, f_2$ nefiind definite ~~pe~~ $R^n \xrightarrow{pe} R^n$ (de exemplu cazul standard) o teoremă care ca 5 (Concluzia 3) nu se poate aplica. S-a recurs atunci la T6. Calculul marginilor (163) oferă însă, o altă posibilitate:

A) Se modifică funcțiile $f_1, f_2$ în afara unui interval suficient de depărtat $(-\infty, M)$ cu $M$ foarte mare (v fig 42)

B) se aplică procedura de aflare a marginilor pentru noile funcții $\tilde{f}_1, \tilde{f}_2$.

C) Dacă marginile se află în domeniul unificat de definiții (verificarea de coincidență), atunci se poate valida calculul lor, înlocuind cu $\tilde{f}_1, \tilde{f}_2$ și orice concluzie privind $\tilde{f}_1, \tilde{f}_2$ poate fi acum aplicată, rezultatele fiind valabile și pentru ecuația originală

(se pot face tot felul de combinații - exemple din concluziile anterioare.)

Mai mult, să observăm că pentru cazurile practice, nu ne interesează soluţiile valori $v_1, v_2$ oricât de mari, (ba chiar ele sînt inacceptabile). De aceea, dacă practica ne impune o limitare de genul

$$(265)\quad \left|\begin{array}{l} v_1 < V_1 \\ v_2 < V_2 \end{array}\right.$$

, putem alege prelungirile $\bar{t}_1, \bar{t}_2$ din fig 42, începînd din $V_1$ şi $V_2$, sau chiar puţin după ele $(V_1+1, V_2+1)$ pentru a îngloba şi eventualele erori: Obţinem:

$\boxed{\text{Consecinţa 10}}$

<u>Dacă</u> $\forall$ <u>$v_1 < V_1$, $v_2 < V_2$</u> impuse de practică, atunci alegînd

$$(266)\quad \bar{t}_1 = \begin{cases} t_1(v_1) & \text{pe } v_1 \le V_1 \\ t_1'(V_1)(v_1 - V_1) + t_1(V_1) & \text{pe } v_1 > V_1 \text{ şi} \end{cases}$$

$$\bar{t}_2 = \begin{cases} t_2(v_2) & \text{pe } v_2 < V_2 \\ t_2'(V_2)(v_2 - V_2) + t_2(V_2) & \text{pe } v_2 > V_2 \end{cases}$$

acestor funcţii li se aplică algoritmul de la Consec. 3 (pag 180) – tocmai unic posibil – şi calculez $\hat{v}_1, \hat{v}_2$. <u>Atunci</u>:

1) Dacă $\hat{v}_1 < V_1$, $\hat{v}_2 < V_2$, atunci ele sînt soluţiile valabile şi pentru ecuaţia originală (eventual iar şi nu e pentru a ține cont de erori)

2) Dacă $\hat{v}_1 > V_1$ sau $\hat{v}_2 > V_2$, atunci ecuaţia originală <u>nu are o soluţie acceptabilă</u>.

---

$\boxed{\text{Consecinţa 11}}$ Dacă se respectă condiţia:

$\sigma_3 \geq \frac{\alpha r}{2(1-\alpha r)}(\sigma_1+\sigma_2)$, putem aplica funcţiilor prelungite algoritmul de la pag 164 (Corolarul 1), urmat de o verificare de coincidenţă:

a) cu impunere practică şi soluţie similară celei precedente

b) Fără impunere. Dacă verificarea nu reuşeşte, funcţiile sînt din nou modificate ($\beta$ mai mare în fig 42). În lucrarea [O] se demonstrează că acest proces se termină după un număr finit de paşi..

## 7) Viteza de convergenţă a algoritmilor de calcul

În paragrafele anterioare, s-au făcut referiri la acest aspect, nu lipsit de importanţă. Problema este că: demonstrarea convergenţei unui algoritm, nu spune nimic despre viteza cu care se produce această convergenţă ("rata de convergenţă")

De exemplu, un algoritm care să se apropie la 0.001 de soluţie la 1000 000 de paşi... Rata lui de convergenţă va fi în acest caz foarte aproape de 1, dar totuşi, subunitară:

(267) $\quad \|x_n - x_*\| \leq \varepsilon^n \|x_0 - x\|$ cu $\varepsilon$ - rata de conv.

Ori pentru noi, o astfel de convergenţă nu are nici o utilitate practică; nu vom încerca nici măcar un calculator electronic cu o astfel de problemă.

Să luăm ca exemplu „algoritmul":

$$(268) \qquad x_{n+1} = \left(\frac{1}{1000}\right)^{\frac{1}{1000\,000}} \cdot x_n, \quad x_0 = 1$$

(presupunem că nu știm că soluția este 0)

Desigur că algoritmul este convergent, căci

$$K = \left(\frac{1}{1000}\right)^{\frac{1}{1000\,000}} < 1$$

și de aceea $x_{n+1} = K x_n$ cu $K < 1$

$$\Rightarrow x_{n+1} = K^n x_0 \to 0$$

(Astfel $\|Tx, Ty\| = |Kx, Ky| = K|x, y|$ $(K < 1)$

și se aplică Th. lui Banach)

Dar, după de-abia 1000 000 pași vom obține

$$x_{1000\,000} = \left[\left(\frac{1}{1000}\right)^{\frac{1}{1000\,000}}\right]^{1000\,000} x_0 = \frac{x_0}{1000} = 0.001$$

Și pentru o precizie dorită de 3 zecimale:

$$|x_n - 0| \le \frac{1}{1000} \text{ numai pentru } n \geq 1000\,000!$$

Desigur că exemplul de mai sus este oarecum particular, dar, el atrage atenția asupra importanței asigurării unei viteze cît mai mari de conver gență a unui algoritm

Pe calea acestui deziderat, să reamintim observația de la pag 189 (U7, lig 42 – f2 (U1)), care ne arată cum putem folosi procedeul de „prelungire", după o prealabilă aflare a marginilor soluțiilor, pentru a „accelera" convergența algoritmilor.

Direcția care urmează (v. art [7]) are mai multe scopuri:

– să exemplifice modele în care se pot face estimări, asupra vitezei de convergență

– să facă o comparație între o metodă de „tip „punct fix" și una de tip „Newton"

– să reflecte importanța condiției de mărginire a părții în teoremele precedente și influența ei asupra vitezei de convergență. Astfel se va înțelege mecanismul prin care procedeele de prelungire pot mări viteza de convergență a unui algoritm.

Se va porni de la o teoremă de punct fix a lui Sandberg, mai generală decât cea de la pag 159 (T2), și se va desemna modul cum poate fi ea particularizată:

**Teorema 11** (Sandberg)

Fie operatorul $G(x) = F(x) + Ax$. Dacă sunt respectate condițiile: $\exists\, k_1, k_2$ cte pozitive ai

$$\begin{cases} \langle G(x) - G(y), x - y \rangle \geq k_1 \|x - y\|^2 & (269) \\ \|G(x) - G(y)\|^2 \leq k_2 \|x - y\|^2 & (270) \end{cases}$$

$\forall x, y \in R^n$ ( $\| \ \|$ norma euclidiană, $\langle \ \rangle$ produs scalar)

Atunci ecuația (1) are o soluție unică și ea este obținută prin trecerea la limită a șirului:

$$x^{k+1} = \frac{k_1}{k_2}\left[B - G(x^k)\right] + x^k \quad \text{pt } k = 1, 2 \ldots \qquad (271)$$

(cu $x^1$ arbitrar)

---

Aplicație

Pentru exemplu de practică, ar trebui verificate condițiile:

$$G(x) = F(x) + Ax = \begin{bmatrix} f_1(x_1) + a_{11}x_1 + a_{12}x_2 \\ f_2(x_1) + a_{21}x_1 + a_{22}x_2 \end{bmatrix} \qquad (272)$$

(și la fel pentru $G(y)$) Deci:

$$G(x) - G(y) = \begin{bmatrix} f_1(x_1) - f_1(y_1) + a_{11}(x_1 - y_1) + a_{12}(x_2 - y_2) \\ f_2(x_2) - f_2(y_2) + a_{21}(x_1 - y_1) + a_{22}(x_2 - y_2) \end{bmatrix} \qquad (273)$$

Sau cu:

$$\frac{f_1(x_1) - f_1(y_1)}{x_1 - y_1} = \alpha_1 > 0 \qquad \frac{f_2(x_2) - f_2(y_2)}{x_2 - y_2} = \alpha_2 > 0 \qquad (274')$$

(unde s-a folosit strict monotonia a lui $k_1, k_2$)

$$G(x)-G(y)=\begin{bmatrix} d_1(x_1-y_1)+a_{11}(x_1-y_1)+a_{12}(x_2-y_2) \\ d_2(x_2-y_2)+a_{21}(x_1-y_1)+a_{22}(x_2-y_2)\end{bmatrix}=$$

$$=\begin{pmatrix} d_1 & 0 \\ 0 & d_2\end{pmatrix}\begin{pmatrix} x_1-y_1 \\ x_2-y_2\end{pmatrix}+A\begin{pmatrix} x_1-y_1 \\ x_2-y_2\end{pmatrix}=\Big\{\overbrace{\Big[\begin{pmatrix} d_1 & 0 \\ 0 & d_2\end{pmatrix}+A\Big]}^{M}\begin{pmatrix} x_1-y_1 \\ x_2-y_2\end{pmatrix}\Big\}$$

$$d_1, d_2 > 0. \qquad (274)$$

Matricea $M$ a fost arătată nesingulară ($A \in P_0$), deci $G(x)-G(y) \neq \Theta$ pt $x \neq y$ și relațiile (269), (270) sunt posibile.

$$E=\langle G(x)-G(y), x-y\rangle=\begin{pmatrix} d_1+a_{11} & a_{12} \\ a_{21} & d_2+a_{22}\end{pmatrix}\begin{pmatrix} x_1-y_1 \\ x_2-y_2\end{pmatrix}\cdot\begin{pmatrix} x_1-y_1 \\ x_2-y_2\end{pmatrix}^T=$$

$$=\begin{bmatrix}(d_1+a_{11})(x_1-y_1)+a_{12}(x_2-y_2) \\ a_{21}(x_1-y_1)+(a_{22}+d_2)(x_2-y_2)\end{bmatrix}\cdot\left[x_1-y_1, x_2-y_2\right]=$$

$$=d_1(x_1-y_1)^2+d_2(x_2-y_2)^2+a_{11}(x_1-y_1)^2+a_{22}(x_2-y_2)^2+$$
$$+(a_{12}+a_{21})(x_1-y_1)(x_2-y_2) \qquad (275)$$

Dorim (rel 269) ca expresia (275) să permită o majorare de genul: $E \geq k_1\|x-y\|^2 = k_1\left[(x_1-y_1)^2+(x_2-y_2)^2\right]$ (276)

Dacă matricea $A$ (la noi $T^{-1}G$) ar fi simetrică discuția s-ar face simplu, cu condiția lui Jordan pentru formele pătratice pozitiv definite. Nu ne aflăm însă în această situație. Va trebui deci să asigurăm (276). O manieră ar fi, de exemplu (de altfel e una uzuală), să cerem ca:

$$(277)\quad d_1(x_1-y_1)^2+d_2(x_2-y_2)^2 \geq k_A\left[(x_1-y_1)^2+(x_2-y_2)^2\right]$$

și

$$(278)\quad a_{11}(x_1-y_1)^2+(a_{12}+a_{21})(x_1-y_1)(x_2-y_2)+a_{22}(x_2-y_2)^2\geq 0$$

Să ne oprim asupra relației (277). <u>Ea ne arată cum apare condiția de pantă mărginită inferior</u> (vezi de ex T 2 pag 159). Într-adevăr, $d_1$ și $d_2$ fiind cei din relația (274'), și condiția 277 va fi asigurată, dacă există margini inferioare strict pozitive $K_A'$, $K_A''$ pentru expresiile (274') deci

$$\frac{f_1(x_1)-f_1(y_1)}{x_1-y_1}\geq K_1'>0 \quad (279)\quad \frac{f_2(x_2)-f_2(y_2)}{x_2-y_2}\geq K_A''>0$$

Putem deci alege

$$\left.\begin{aligned} K_1' &= \inf\left\{\frac{f_1(x_1)-f_1(y_1)}{x_1-y_1}\right\}\\ K_A'' &= \inf\left\{\frac{f_2(x_2)-f_2(y_2)}{x_2-y_2}\right\}\\ &\boxed{K_A=\min\{K_1',K_A''\}}\end{aligned}\right|\quad(280)$$

și condiția 277 este respectată

Să ne oprim asupra condiției (278). Ea este echivalentă cu:

(281) $a_{11}u^2 + (a_{12}+a_{21})uv + a_{22}v^2 \geq 0 \quad \forall u,v$

Putem proceda direct, variind aceasta ca un polinom în $u$, punem condiția ca

(282') $\Delta < 0$ (stiu că $a_{11} > 0$)

(282) $(a_{12}+a_{21})^2 v^2 - 4a_{11}a_{22}v^2 \leq 0 \quad \forall v$ sau

(283) $(a_{12}+a_{21})^2 \leq 4a_{11}a_{22}$

Astfel (și mai general pe $n$ mai mare), scriem expresia (281) în forma:

$$(u\ v)\begin{pmatrix} a_{11} & \frac{a_{12}+a_{21}}{2} \\ \frac{a_{12}+a_{21}}{2} & a_{22}\end{pmatrix}\begin{pmatrix} u \\ v\end{pmatrix} \geq 0 \quad \forall (u,v)$$

și acum putem aplica t. pentru forme pătratice

$a_{11} \geq 0$

$$D = \begin{vmatrix} a_{11} & \frac{a_{12}+a_{21}}{2} \\ \frac{a_{12}+a_{21}}{2} & a_{22}\end{vmatrix} \geq 0 \text{ și regăsim } 283$$

Acest procedeu poate fi aplicat în general (ruperea matricii $A$ în $\begin{cases} A_1 \text{ simetrică} \\ A_2 \text{ antisimetrică} \quad uA^2u^T = 0\end{cases}$)

Pentru problema noastră condiția este:

$$K^2\left[G_3(\alpha_t-1) + G_3(\alpha_r-1) + \alpha_r(G_1+G_2)\right]^2 \leq \frac{4}{K^2}\left[G_3(1-\alpha_r)(G_1+G_2+G_3(1-\alpha_t)\right]$$

(v. expr. lui $A_1$ de la 206 pag 192)

$$\left[G_3(\alpha_r+\alpha_t-2)+\alpha_r(G_1+G_2)\right]^2 \le 4G_3(1-\alpha_r)\left(G_1+G_2+G_3(1-\alpha_t)\right)$$

$$G_3^2(\alpha_r+\alpha_t-2)^2+\alpha_r^2(G_1+G_2)^2-2\alpha_r G_3(G_1+G_2)(2-\alpha_t-\alpha_r)$$

$$\le 4G_3(1-\alpha_r)(G_1+G_2)+4G_3^2(1-\alpha_r)(1-\alpha_t)$$

sau

$$A\,G_3^2+B\,(G_1+G_2)^2+C\,G_3\,(G_1+G_2)\le 0 \qquad (284)$$

unde $A=(\alpha_r+\alpha_t-2)^2-4(1-\alpha_r)(1-\alpha_t)=$

$=(u+v)^2-4uv=(u-v)^2\geq 0.$

$B=\alpha_r^2\geq 0$

$C=-\left[2\alpha_r(2-\alpha_t-\alpha_r)+4(1-\alpha_r)\right]=-\left[\cancel{4\alpha_r}-2\alpha_r\alpha_t+2\alpha_r^2+4\cancel{-4\alpha_r}\right]$
$<0.$

Am reluat acest calcul pentru a sublinia că nu putem avea certitudinea dorită decât într-o gamă restrânsă a parametrilor, căci $A>0 \Rightarrow (282)\to\infty$ pentru $G_3\to\infty$ iar $C>0 \Rightarrow (282)\to\infty$ pentru $(G_1+G_2)\to\infty$, și deci într-un caz nu se mai poate respecta condiția de $\le 0$.

Pentru un caz tipic $\alpha_r \sim 0{,}98 \sim 1$

$\alpha_t \sim 0{,}56 \sim 0{,}5$

$\Rightarrow A=(1-\alpha_r+\alpha_t-1)^2=(\alpha_r-\alpha_t)^2=0{,}5^2=0{,}25$

$B = \alpha_t^2 = 0.25$

$C = -\left[4 + 2\alpha_r^2 - 2\alpha_r \alpha_t\right] = -\left[4 + 2\cdot 0.25 - 2\cdot 0.5\right] = 3{,}5$

Deci condiția este:

$$0.25\, G_3^2 + 0.25 (G_2 + G_1)^2 - 3{,}5\, G_3 (G_1 + G_2) \leq 0. \quad \text{sau}$$

$$\left| G_3^2 + (G_1+G_2)^2 - 10\, G_3 (G_1+G_2) \right) \leq 0 \qquad (285)$$

Dar: $\left| G_3^2 + (G_1+G_2)^2 - 2\left(G_3(G_1+G_2)\right) \right| \geq 0$ și de aici

se vede gama relativ restrânsă în care trebuie să ne încadrăm. (Calculele mai exacte vor plasa pe $(G_1+G_2)$ între rădăcinile polinomului, variind cu variabila $G_3$).

Aceasta sugerează posibilitatea „surprizei" pe care o poate oferi calculatorul echipat cu acest program, pentru anumite valori ale parametrilor!

Urmărind cu totul analog îndeplinirea condiției (270) se va ajunge la o nouă dublă condiționare. Pe noi ne interesează condiția asupra punții. Ea va fi evident de sens invers:

$$\left\{\begin{array}{l} K_2' = \sup\left\{ \dfrac{t_1(x_1) - t_1(y_1)}{x_1 - y_1} \right\} \\ K_2'' = \sup\left\{ \dfrac{t_2(x_2) - t_2(y_2)}{x_2 - y_2} \right\} \qquad (286) \\ K_2 = \max\left\{ K_2', K_2'' \right\} \end{array}\right.$$

Astfel, am încercat să sugerez doar, modul cum apar condițiile dintr-o teoremă clasică a lui Sandberg ( $A$ pozitiv semidefinită, iar $f_K$ cu pantele mărginite inferior și superior).

Deasemenea, observăm, din relațiile (280)(286) că numerele $K_1$, $K_2$ pot fi alese în funcție de minimul sau maximul pantelor. Astfel:

inf pantei mare $\Rightarrow K_1$ mare

sup pantei mic $\Rightarrow K_2$ mic

Or, în demonstrația sa, Sandberg arată că rata convergenței ( $K$-ul contracției pe care o construiește) este ~~proporțională~~ dependentă de $K = f\left(\frac{K_1}{K_2}\right)$. Desigur că $K_f$-ul e arătat asemănător (contracție), dar rămâne deschisă problema vitezei de convergență, mai ales pentru $K_1$ foarte mici și $K_2$ foarte mari, caz în care $K \to 1$.

Pentru a intra în aceste detalii și a înțelege rolul procesului de modificare din fig 42, care în fond mărește pe "inf pantei" (și deci pe $K_1$ și deci viteza de convergență) să urmărim un raționament făcut de Wilson (vezi [7])

El a observat că, pornind de la relațiile (269) și (270) și aplicând ineg. clasică a lui Schwartz:

$\langle a, b\rangle^2 \leq \|a\|\,\|b\|$ (287) se obțin două inegalități cu sens schimbat:

$$\langle G(x)-G(y), x-y\rangle^2 \leq \|G(x)-G(y)\|\,\|x-y\| \overset{(270)}{\leq} k_2\|x-y\|^2$$

de unde:

$$\langle G(x)-G(y), x-y\rangle \leq \underset{K_3}{\sqrt{k_2}}\,\|x-y\| \quad (288)$$

$$\|G(x)-G(y)\|^2\,\|x-y\|^2 \geq \langle (G(x)-G(y)), x-y\rangle^4 \overset{(269)}{\geq}$$

$$\geq k_1^2\|x-y\|^4 \quad \text{deci}$$

$$\|G(x)-G(y)\|^2 \geq \underset{K_4}{k_1^2}\,\|x-y\|^2 \quad (289)$$

Această inversare îi permite să urmărească toată demonstrația lui Sandberg cu sensul schimbat al inegalităților și $K_3, K_4$ în loc de $K_1, K_2$, ajungând la:

$$\|H(x)-H(y)\| \geq K\,\|x-y\|^2 \quad (290)$$

($H(x)$ funcția contracție)

cu $K = 1 - 2\left(\frac{K_4}{K_3^2}\right)^{1/2} + \left(\frac{K_4}{K_3^2}\right)^2$ sau

$$K = 1 - 2\left(\frac{k_1^2}{k_2}\right)^{1/2} + \left(\frac{k_1^2}{k_2}\right)^2 \quad (291)$$

(în teoremă

$\Rightarrow \|H(x)-H(y)\| \leq \bar{K}\,\|x-y\|^2$ cu $\bar{K} < 1$)

(292)

Deci $x_0$ fiind punct fix al contracției, din (290) se deduce imediat

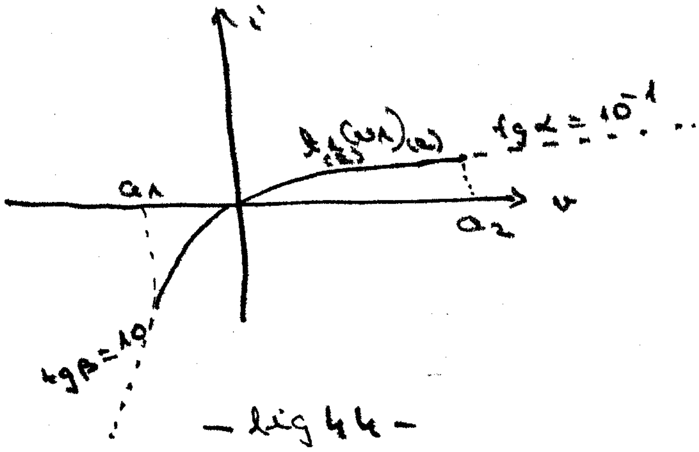
$$\|x^{k+1}-x_0\| = \|H(x^k)-H(x^0)\| \geq K\,\|x^k-x_0\| \quad (293)$$

Această inegalitate este o foarte interesantă, care oferă o limită inferioară a ratei de convergență ($\sqrt{K}$). (din 291). Nelson arată mai întâi că $K>0$ deci (293) nu este trivială. Mai mult încă, dacă $K_1$ devine foarte mic iar $K_2$ foarte mare, $\sqrt{K}\to 1$ și deci șirul devine foarte lent convergent

Exemplu

Dacă eroarea este să nu putem mărgini pantele funcțiilor prea mult, adică ni se impune de exemplu

$$10^{-1} < \frac{f_i(x)-f_i(y)}{x-y} < 10^{1} \quad (294)$$




De exemplu atunci când nu putem realiza mărgini ale soluțiilor mai bune ca $a_1, a_2$

— fig 44 —

$$K = 1 - 2\left(\frac{10^{-2}}{10}\right)^{1/2} + \left(\frac{10^{-2}}{10}\right)^{2} \qquad \boxed{\sqrt{K}\approx 0.99} \quad (295)$$

$\Rightarrow$ Concluzie: indiferent de numărul iterațiilor se va câștiga în mod cert mai mult de 1% de

<u>apropierea de soluție, la un pas</u> (cu $K_1$ mai mic și $K_2$ mai mare situația se poate agrava. Totuși, de obicei nu putem admite funcționări în regiuni foarte depărtate).

În schimb, Nelson arată că schema lui (T4 <u>pag 168</u>), deși mai puțin generală (v. condițiile) atunci când se poate aplica, va ar putea converge mult mai repede la soluție. (dacă funcțiile sunt liniarizate pe porțiuni, numărul de pași va fi chiar finit)

Să rescriem deci rel. (157):

$$x^{k+1} = [F'(x^k) + A]^{-1} [B - F(x^k) + F'(x^k) x^k] \quad (157')$$

$$x^0 = [F'(x^k) + A]^{-1} [B - F(x_0) + F'(x^k) x^0] \quad (296)$$

(această identitate se verifică ușor, ținînd cont că $x^0$ este soluția ecuației $F(x) + Ax = B$)

Scăzînd relațiile de mai sus obținem:

$$x^{k+1} - x^0 = [F'(x^k) + A]^{-1} [F(x^0) - F(x^k) - F'(x^k)(x^0 - x^k)] \quad (297)$$

relație interesantă, care relevă faptul că, pe măsură ce ne apropiem de soluție, caz în care:

$$\frac{f_i(x_i^0) - f_i(x_i^k)}{x_i^0 - x_i^k} \approx f_i'(x_i^k) \quad (298)$$

deci $F(x^0) - F(x^k) \approx F'(x^0)(x^0 - x_k)$ (299)

se este evident că, în relația (297), $\|x^{k+1} - x^0\|$ este tot mai mic, chiar dacă $\|x^k - x^0\|$ este încă suficient de mare: La fiecare pas, apropierea este mai puternică (rata convergenței devine nelimitată, pe măsura apropierii de soluție) - o „accelerare" a convergenței.

Aș dori ca acum să apară mai clar unitatea teoremelor de analiză calitativă fundamentală (A.C.F), înlănțuirea strânsă între problemele existenței, unicității, mărginirilor, continuității soluției, a convergenței și bunei definiri a algoritmilor de calcul.