Μάθημα 1ο Ε3. Υπολογιστικά Μαθηματικά Ι 1/10/2018

Διαλέξεις: Δευτέρα-Τετάρτη: 9-11
Βιβλίο: Θα δοθούν σημειώσεις του Β. Δουγαλή
Ασκήσεις: Θα δοθούν σετ θεωρητικών ασκήσεων (1 σε κάθε ενότητα)
και 1 υπολογιστική άσκηση, μπορεί να γίνει σε Matlab ή σε κάποια γλώσσα προγραμματισμού

Περιεχόμενο του μαθήματος

1. Αριθμητική λύση μη γραμμικών συστημάτων
2. Αριθμητική Γραμμική Άλγεβρα (Απαλοιφή Gauss και επαναληπτικές μέθοδοι)
3. Παρεμβολή (Κλασσική πολυωνυμική παρεμβολή, και με Splines)
4. Αριθμητική Ολοκλήρωση.

Βιβλιογραφία

1. Γ. Ακρίβης - Β. Δουγαλής, "Εισαγωγή στην Αριθμητική Ανάλυση", Π.Ε.Κ. 4η έκδοση 2010
2. R.L. Burden, J.D. Faires and A.M. Burden, "Numerical Analysis" (Brooks/Cole) 10^{th} edition 2015
3. W. Gautschi, "Numerical Analysis, An introduction", Birkhauser, 2^{nd} edition, 2012
4. J. Stoer, R. Bulirsch, "Introduction to Numerical Analysis", Springer 3^{rd} edition, 2002.

Νόρμες διανυσμάτων και πινάκων

Ορισμός: Μια διανυσματική νόρμα στον $\mathbb{R}^n$ είναι μια απεικόνιση $\|\cdot\|:\mathbb{R}^n\to\mathbb{R}$, $x\to\|x\|$ με τις ιδιότητες:

(i) $\|x\|\geq 0$, $\forall x\in\mathbb{R}^n$ και $\|x\|=0 \Rightarrow x=0$

(ii) $\|\lambda x\| = |\lambda|\cdot\|x\|$, $\forall x\in\mathbb{R}^n, \lambda\in\mathbb{R}$

(iii) $\|x+y\|\leq\|x\|+\|y\|$ $\forall x,y\in\mathbb{R}^n$

Μερικές νόρμες που χρησιμοποιούνται στην Αριθμητική Ανάλυση είναι:

$\|x\|_\infty = \max\limits_{1\leq i\leq n}|x_i|$, $x=(x_1,x_2,\dots,x_n)^T$ (max νόρμα)

$\|x\|_1 = \sum\limits_{i=1}^{n}|x_i|$

$\|x\|_2 = \left(\sum\limits_{i=1}^{n}|x_i|^2\right)^{1/2}$ (Ευκλείδια νόρμα)

Για την Ευκλείδια νόρμα ισχύει $\|\cdot\|_2 = \sqrt{(\cdot,\cdot)_2}$, όπου $(x,y)_2 = \sum\limits_{i=1}^{n} x_i y_i$, $x,y\in\mathbb{R}^n$

Σημαντικό είναι ότι στον $\mathbb{R}^n$ ισχύει η ισοδυναμία των νορμών.

Ορισμός: Δύο νόρμες $\|\ \|_\alpha$ και $\|\ \|_\beta$ στον $\mathbb{R}^n$ είναι ισοδύναμες (ή συγκρίσιμες) αν υπάρχουν θετικές σταθερές C_1 και C_2 (που εξαρτώνται από τις νόρμες α και β) τέτοιες ώστε:

$$C_1 \cdot \|x\|_\alpha \le \|x\|_\beta \le C_2 \|x\|_\alpha \quad \forall x \in \mathbb{R}^n$$

Πρόταση: Όλες οι νόρμες στον $\mathbb{R}^n$ είναι ισοδύναμες μεταξύ τους.

Απόδειξη: Βιβλίο Ακρίβη-Δουγαλή πρόταση 3.1.

Ορισμός: Λέμε ότι η ακολουθία $\{x^k\}_{k=1}^{\infty}$ διανυσμάτων του $\mathbb{R}^n$ συγκλίνει στο διάνυσμα $x \in \mathbb{R}^n$ (γράφουμε $x^k \to x$ καθώς $k \to \infty$) αν για κάποια νόρμα $\|\cdot\|$ του $\mathbb{R}^n$ ισχύει $\lim_{k\to\infty} \|x^k - x\| = 0$.

Είναι προφανές ότι λόγω της ισοδυναμίας των νορμών στον $\mathbb{R}^n$, η σύγκλιση της ακολουθίας είναι ανεξάρτητη της επιλεγόμενης νόρμας.

Επίσης, λόγω της ισοδυναμίας της $\|\cdot\|$ με την $\|\cdot\|_\infty$ έπεται ότι

$$x^k \to x \text{ καθώς } k \to \infty \iff \forall i,\ 1 \le i \le n \quad x_i^k \to x_i \text{ καθώς } k \to \infty$$

Κάθε $\|\cdot\|$ του $\mathbb{R}^n$ παράγει μια αντίστοιχη νόρμα στον $\mathbb{R}^{n\times n}$

Ορισμός: Έστω $\|\cdot\|$ μια νόρμα του $\mathbb{R}^n$. Η απεικόνιση $\|\cdot\|: \mathbb{R}^{n\times n} \longrightarrow \mathbb{R}$

$$\|A\| = \sup_{\substack{x\in\mathbb{R}^n \\ \|x\|\le 1}} \|Ax\| \quad \left(= \sup_{\substack{x\in\mathbb{R}^n \\ \|x\| = 1}} \|Ax\| = \sup_{\substack{x\in\mathbb{R}^n \\ x\neq 0}} \frac{\|Ax\|}{\|x\|} \right)$$

Η παραπάνω απεικόνιση είναι καλώς ορισμένη.

Έστω $x = (x_i)_{i=1}^{n\,T} \in \mathbb{R}^n$, $\|x\| \le 1$ και $A = (\alpha_{ij})_{i,j=1}^{n}$

Τότε λόγω της ισοδυναμίας των νορμών $\|\ \|$ και $\|\cdot\|_\infty$ έχουμε το εξής:

$$\|Ax\| \le C_1 \|Ax\|_\infty = C_1 \max_{1\le i\le n} \left|\sum_{j=1}^{n} \alpha_{ij} x_j\right| \le C_1 \max_{1\le i\le n} \sum_{j=1}^{n} |\alpha_{ij}|\cdot|x_j|$$

$$\le C_1 \max_{1\le i\le n} |x_i| \max_{1\le i\le n} \sum_{j=1}^{n} |\alpha_{ij}| \le C_1 \cdot \|x\|_\infty \max_{1\le i\le n} \sum_{j=1}^{n} |\alpha_{ij}| \le C_1 C_2 \|x\| \max_{1\le i\le n} \sum_{j=1}^{n} |\alpha_{ij}|$$

$$\overset{\|x\|\le 1}{\Longrightarrow} \le C_1 \cdot C_2 \cdot \max_{1\le i\le n} \sum_{j=1}^{n} |\alpha_{ij}| \quad (\text{ανεξάρτητο του } x).$$

E3. Υπολογιστικά Μαθηματικά Ι 1/10/2018

Επίσης, η απεικόνιση αυτή πληροί τις ιδιότητες:

(i) $\|A\| \geq 0 \quad \forall\ A \in \mathbb{R}^{n\times n}$ και $\|A\| = 0 \Rightarrow A = O$

(ii) $\|\lambda \cdot A\| = |\lambda|\,\|A\| \quad \forall\ A \in \mathbb{R}^{n\times n}, \lambda \in \mathbb{R}$

(iii) $\|A+B\| \leq \|A\| + \|B\| \quad \forall\ A, B \in \mathbb{R}^{n\times n}$

και

(iv) $\|Ax\| \leq \|A\| \cdot \|x\| \quad \forall\ A \in \mathbb{R}^{n\times n}, x \in \mathbb{R}^n$

(v) $\|A \cdot B\| \leq \|A\| \cdot \|B\| \quad \forall\ A, B \in \mathbb{R}^{n\times n}$

(vi) $\|I_n\| = 1$

(για απόδειξη βλέπουμε Ακρίβη-Δουγαλή)

Μάθημα 2ο   Ε3. Υπολογιστικά Μαθηματικά Ι   3/10/2018

Νόρμες διανυσμάτων και πινάκων

Ορισμός: Έστω $\|\cdot\|$ μια νόρμα στον $\mathbb{R}^n$. Η απεικόνιση
$\|\cdot\| : \mathbb{R}^{n\times n} \longrightarrow \mathbb{R}$, $\|A\| = \sup\limits_{\substack{x\in\mathbb{R}^n \\ \|x\|\le 1}} \|Ax\|$ λέγεται φυσική νόρμα πινάκων

Οι νόρμες πινάκων που παράγονται από τις $\|\cdot\|_\infty$, $\|\cdot\|_1$ και $\|\cdot\|_2$ του $\mathbb{R}^n$, αν $A = (\alpha_{ij})_{i,j=1}^{n}$ είναι:

$\|A\|_\infty = \max\limits_{1\le i\le n} \sum\limits_{j=1}^{n} |\alpha_{ij}|$ (νόρμα του αθροίσματος γραμμών)

$\|A\|_1 = \max\limits_{1\le j\le n} \sum\limits_{i=1}^{n} |\alpha_{ij}|$ (νόρμα του αθροίσματος στηλών)

$\|A\|_2 = \max\limits_{1\le i\le n} [\lambda_i(A^T A)]^{1/2}$ (φασματική νόρμα), όπου $(A^T)_{ij} = \alpha_{ji}$, $1\le i,j\le n$ και όπου $\lambda_i(B)$ συμβολίζουμε τις ιδιοτιμές του πίνακα $B\in\mathbb{R}^{n\times n}$

Από τον ορισμό των νορμών πινάκων προκύπτει ότι δύο οποιεσδήποτε νόρμες πινάκων είναι ισοδύναμες μεταξύ τους. Επίσης, ότι ισχύει για την σύγκλιση μιας ακολουθίας διανυσμάτων ισχύει και σε μια ακολουθία πινάκων.

Πρόταση: Έστω $\|\cdot\|$ μια νόρμα του $\mathbb{R}^n$ και $\|\cdot\|$ η νόρμα στον $\mathbb{R}^{n\times n}$ που παράγεται από αυτήν. Αν $A\in\mathbb{R}^{n\times n}$ και $\|A\|\le 1$, τότε ο πίνακας $I_n \pm A$ είναι αντιστρέψιμος και ισχύει:

$$\frac{1}{1+\|A\|} \le \|(I_n \pm A)^{-1}\| \le \frac{1}{1-\|A\|}$$

Απόδειξη: Άσκηση

Πόρισμα: Έστω $\|\cdot\|$ μια νόρμα του $\mathbb{R}^n$ και $\|\cdot\|$ η νόρμα του $\mathbb{R}^{n\times n}$ που παράγεται από αυτήν. Αν $A\in\mathbb{R}^{n\times n}$ αντιστρέψιμος και $B\in\mathbb{R}^{n\times n}$ τέτοιοι ώστε $\|A^{-1}\|\,\|B\|\le 1$, τότε ο πίνακας $A\pm B$ είναι αντιστρέψιμος και επιπλέον ισχύει ότι:

$$\|(A\pm B)^{-1}\| \le \frac{\|A^{-1}\|}{1-\|A^{-1}\|\,\|B\|}$$

Απόδειξη: Θεωρήστε κατάλληλο πίνακα (?) στη θέση του A της προηγούμενης πρότασης και κατόπιν εφαρμόστε την.

Ορισμός: Η φασματική ακτίνα $\rho(A)$ ενός πίνακα A, ορίζεται ως εξής:
$\rho(A) = \max\limits_{1\leq i\leq n} |\lambda_i(A)|$

Με αυτόν τον ορισμό, $\|A\|_2 = [\rho(A^T A)]^{1/2}$
Επίσης, λ_i ιδιοτιμή του A σημαίνει ότι $\exists\, u \in \mathbb{R}^n$, $u \neq 0$, τέτοιο ώστε $A\cdot u = \lambda_i \cdot u$
Τώρα αν $\|\cdot\|$ είναι μια νόρμα του $\mathbb{R}^n$ και $\|\cdot\|$ η αντίστοιχη παραγώμενη στον $\mathbb{R}^{n\times n}$
$\lambda_i u = A u \Rightarrow \|\lambda_i u\| = \|Au\| \Rightarrow |\lambda_i| \|u\| = \|Au\| \leq \|A\| \|u\|$, $\|u\| \neq 0$
$\Rightarrow |\lambda_i| \leq \|A\|$ $i=1,2,\dots,n \Rightarrow \rho(A) \leq \|A\|$

Σχεδόν αντίστροφο αποτέλεσμα
Λήμμα: Έστω $A \in \mathbb{R}^{n\times n}$ τότε $\forall\, \varepsilon > 0\ \exists$ νόρμα $\|\cdot\|$ στον $\mathbb{R}^n$ τέτοια ώστε $\|A\| \leq \rho(A) + \varepsilon$
Απόδειξη: Σημειώσεις Δουγαλή.

1. Αριθμητική λύση μη γραμμικών συστημάτων

Θα ασχοληθούμε με την αριθμητική επίλυση μη γραμμικών συστημάτων n εξισώσεων με n αγνώστους της μορφής:
$f_i(x_1, x_2, \dots, x_n) = 0$ $i=1,2,\dots,n$ ή όπως γράφεται διανυσματικά
$F(x) = 0$, όπου $F: D \subset \mathbb{R}^n \to \mathbb{R}^n$, $F(x) = (f_1(x), f_2(x), \dots, f_n(x))^T$, $x = (x_1, \dots, x_n)^T$

1.1 Παραγωγίσιμες συναρτήσεις στον $\mathbb{R}^n$

Ορισμός 1: Λέμε ότι η απεικόνιση $F: D \subseteq \mathbb{R}^n \to \mathbb{R}^n$ είναι παραγωγίσιμη σ'ένα σημείο $x \in \operatorname{int} D$ (με την έννοια του Frechet) αν υπάρχει γραμμικός τελεστής (ή γραμμικός μετασχηματισμός) $A_x: \mathbb{R}^n \to \mathbb{R}^n$ τέτοιος ώστε για κάποια νόρμα του $\mathbb{R}^n$ να ισχύει:
$$\lim_{h\to 0} \frac{\|F(x+h) - F(x) - A_x h\|}{\|h\|} = 0 \qquad (A_x h \equiv A_x(h))$$

Ε3. Υπολογιστικά Μαθηματικα Ι 3/10/2018

Παρατηρήσεις:

1. Η A_x εξαρτάται από το x (όχι γραμμικά) αλλά δεν εξαρτάται από την επιλογή της νόρμας (λόγω της ισοδυναμίας των νορμών στον $\mathbb{R}^n$)

2. Ο προηγούμενος ορισμός γενικεύει την έννοια της παραγωγισιμότητας στον $\mathbb{R}^1$ αν $f:\mathbb{R}^1 \to \mathbb{R}^1$ η f λέγεται παραγωγίσιμη σε ένα σημείο $x\in\mathbb{R}^1$, αν υπάρχει ένα $\alpha_x \in \mathbb{R}$ τέτοιος ώστε $\lim_{h\to 0}\frac{f(x+h)-f(x)}{h} = \alpha_x \Leftrightarrow \lim_{h\to 0}\frac{|f(x+h)-f(x)-\alpha_x h|}{|h|}=0$
Το α_x εξαρτάται από το x και γράφουμε $\alpha_x = f'(x)$

3. Αν η F είναι παραγωγίσιμη στο x, τότε ο γραμμικός τελεστής A_x είναι μοναδικός. Πράγματι έστω ότι υπάρχουν δυο γραμμικοί τελεστές A_1 και A_2 τέτοιοι ώστε να ισχύει

$$\frac{\|F(x+h)-F(x)-A_1h\|}{\|h\|} \longrightarrow 0 \quad \text{καθώς } h\to 0$$

$$\frac{\|F(x+h)-F(x)-A_2h\|}{\|h\|} \longrightarrow 0 \quad \text{καθώς } h\to 0$$

Τότε $\forall y \in \mathbb{R}^n$, $y\neq 0$ (σταθερό), έχουμε (θέτοντας $h=ty$)

$$\frac{\|F(x+ty)-F(x)-A_1(ty)\|}{\|ty\|} \longrightarrow 0 \text{ καθώς } t\to 0 \text{ και } \frac{\|F(x+ty)-F(x)-A_2(ty)\|}{\|ty\|} \longrightarrow 0, \text{ καθώς } t\to 0$$

Τώρα
$$\frac{\|(A_1-A_2)y\|}{\|y\|} = \frac{|t|\,\|(A_1-A_2)y\|}{|t|\,\|y\|} = \frac{\|t(A_1-A_2)y\|}{\|ty\|} = \frac{\|(A_1-A_2)ty\|}{\|ty\|} =$$

$$= \frac{\|A_1(ty)-A_2(ty)+F(x+ty)-F(x)-F(x+ty)+F(x)\|}{\|ty\|} = \frac{\|[F(x+ty)-F(x)-A_2(ty)]-[F(x+ty)-F(x)-A_1(ty)]\|}{\|ty\|} \leq$$

$$\leq \frac{\|F(x+ty)-F(x)-A_2(ty)\|}{\|ty\|} + \frac{\|F(x+ty)-F(x)-A_1(ty)\|}{\|ty\|} \longrightarrow 0 \quad \text{καθώς } t\to 0$$

Άρα $\forall y\in\mathbb{R}^n$ $y\neq 0$: $\frac{\|(A_1-A_2)y\|}{\|y\|}\leq 0 \implies \|(A_1-A_2)y\|=0 \quad \forall y \in \mathbb{R}^n$

$$\implies (A_1-A_2)y = 0 \quad \forall y\in\mathbb{R}^n \implies A_1-A_2=0 \implies A_1=A_2$$

Ορισμός 2: Αν η F είναι παραγωγίσιμη στο $x\in \operatorname{Int} D$, λέμε ότι ο γραμμικός τελεστής A_x είναι η παράγωγος της F και γράφουμε $A_x = F'(x)$

Μάθημα 3ο Ε3. Υπολογιστικά Μαθηματικά Ι 8/10/2018

Ορισμός 3: Η F είναι παραγωγίσιμη σ' ένα ανοικτό σύνολο $D_0 \subset \mathrm{int}\, D$, αν και μόνο αν η F είναι παραγωγίσιμη στο D_0, τότε η F' μπορεί να θεωρηθεί σαν η απεικόνιση του D_0 στο σύνολο $L(\mathbb{R}^n)$ των γραμμικών τελεστών του $\mathbb{R}^n$ στον εαυτό του.

Παρατήρηση: Μπορούμε εύκολα να δείξουμε ότι: Αν $F_1, F_2 : D \subset \mathbb{R}^n \longrightarrow \mathbb{R}^n$ είναι παραγωγίσιμες στο $x \in \mathrm{Int}D$, τότε το ίδιο ισχύει για την $\alpha \cdot F_1 + \beta \cdot F_2$, $\alpha, \beta \in \mathbb{R}$ και μάλιστα $(\alpha \cdot F_1 + \beta F_2)'(x) = \alpha \cdot F_1'(x) + \beta\, F_2'(x)$. (Αφήνεται σαν άσκηση)

Δηλαδή η πράξη της παραγώγισης διέπεται από γραμμικότητα

Παρένθεση σε όσα λέμε:

Αν X είναι ένας γραμμικός χώρος, ο $A : X \longrightarrow X$ γραμμικός τελεστής και $\{x_1, x_2, \ldots, x_n\}$ μια βάση του X, τότε υπάρχουν πραγματικοί αριθμοί α_{ij} τέτοιοι ώστε: $Ax_j = \sum_{i=1}^{n} \alpha_{ij} x_i$, $1 \le j \le n$, έτσι αν θέσουμε:

$$[A] = \begin{bmatrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \cdots & \alpha_{2n} \\ \vdots & \vdots & \ddots & \\ \alpha_{n1} & \alpha_{n2} & \cdots & \alpha_{nn} \end{bmatrix}$$

έχουμε για παράδειγμα: $Ax_2 = \alpha_{12} x_1 + \alpha_{22} x_2 + \ldots + \alpha_{n2} x_2$, δηλαδή το πεδίο τιμών του γραμμικού τελεστή A παράγεται από τις στήλες του $[A]$.

Τώρα αν $x \in X$, $x = \sum_{j=1}^{n} c_j x_j$ και χρησιμοποιώντας την γραμμικότητα του τελεστή A, έχουμε:

$$Ax = A\left(\sum_{j=1}^{n} c_j x_j\right) = \sum_{j=1}^{n} c_j Ax_j = \sum_{j=1}^{n} c_j \left(\sum_{i=1}^{n} \alpha_{ij} x_i\right) = \sum_{i=1}^{n} \left(\sum_{j=1}^{n} \alpha_{ij} c_j\right) x_i,$$

δηλαδή οι συντεταγμένες του Ax είναι $\sum_{j=1}^{n} \alpha_{ij} c_j$

Αντιστρόφως δοθέντος ενός πραγματικού πίνακα $[A]$ με στοιχεία α_{ij}, αν ορίσουμε $A : X \to X$, $Ax = \sum_{i=1}^{n} \left(\sum_{j=1}^{n} \alpha_{ij} c_j\right) x_i$ τότε $A \in L(X)$

Επομένως, υπάρχει μια 1-1 αντιστοιχία ανάμεσα στον $L(X)$ και τον $\mathbb{R}^{n \times n}$

Παρατήρηση: Ο πίνακας [A] εξαρτάται, όχι μόνο από τον A, αλλά και από την επιλογή της βάσης του X. Διαφορετικές βάσεις δίνουν διαφορετικούς πίνακες [A] για τον ίδιο γραμμικό τελεστή A

Ερώτημα: Πως παριστάνεται (υπό μορφή πίνακα) η $F'(x)$;

Αν $\{e_1, e_2, \dots, e_n\}$ είναι κανονική βάση του $\mathbb{R}^n$, τότε $F(x) = \sum_{i=1}^{n} f_i(x) e_i$, $x \in D$

Για $x \in D$, η μερική παράγωγος $\frac{\partial f_i}{\partial x_j}$ της f_i ως προς x_j ορίζεται ως εξής:

$\frac{\partial f_i}{\partial x_j}(x) = \lim_{t \to 0} \frac{f_i(x + te_j) - f_i(x)}{t}$, $1 \le i, j \le n$ υπό την προϋπόθεση ότι το όριο $\exists$.

Πρόταση 1: Έστω $F: D \subset \mathbb{R}^n \to \mathbb{R}^n$, $x \in \mathrm{Int}\, D$ και F παραγωγίσιμη στο x. Τότε οι μερικές παράγωγοι $\frac{\partial f_i}{\partial x_j}(x)$, $1 \le i, j \le n$ υπάρχουν και $F'(x) e_j = \sum_{i=1}^{n} \frac{\partial f_i}{\partial x_j}(x) e_j$, $1 \le j \le n$

Απόδειξη: Εφόσον η F είναι παραγωγίσιμη στο X, $\lim_{h \to 0} \frac{\|F(x+h) - F(x) - F'(x)h\|}{\|h\|} = 0$

Σταθεροποιώντας το j και θέτοντας $h = t \cdot e_j$, οπότε $h \to 0 \Leftrightarrow t \to 0$ έχουμε με $\|\cdot\| = \|\ \|_\infty$,

$$\lim_{t \to 0} \frac{\|F(x+te_j) - F(x) - F'(x)(te_j)\|_\infty}{\|te_j\|_\infty} = 0 \xrightarrow[\|te_j\| = |t|\|e_j\|_\infty = |t|]{\text{Γραμμικότητα } F'(x)} \lim_{t \to 0} \frac{\|F(x+te_j) - F(x) - tF'(x)e_j\|_\infty}{|t|} = 0 \quad (1)$$

Δεδομένου ότι $F(x) = \sum_{i=1}^{n} f_i(x) e_j$ και $F'(x) e_j = \sum_{i=1}^{n} F'(x)_{ij} e_i$, $F'(x)_{ij} = [F'(x)]_{ij}$,

$F(x+te_j) - F(x) - t\,F'(x)e_j = \sum_{i=1}^{n} [f_i(x+te_j) - f(x) - t\,F'(x)_{ij}] e_i$ οπότε η (1) δίνει

$$\lim_{t \to 0} \frac{|f_i(x+te_j) - f_i(x) - t\,F'(x)_{ij}|}{|t|} = 0 \Leftrightarrow \lim_{t \to 0} \left| \frac{f_i(x+te_j) - f_i(x)}{t} - F(x)_{ij} \right| = 0 \quad 1 \le i, j \le n,$$

$$\Leftrightarrow \lim_{t \to 0} \frac{f_i(x+te_j) - f_i(x)}{t} = F'(x)_{ij} \Leftrightarrow \frac{\partial f_i}{\partial x_j} = F'(x)_{ij} \quad 1 \le i, j \le n$$

Πρόταση 2: Έστω $F: D \subset \mathbb{R}^n \to \mathbb{R}^n$ και $x \in \mathrm{Int}\, D$. Αν οι μερικές παράγωγοι $\frac{\partial f_i(x)}{\partial x_j}$, $1 \le i, j \le n$ υπάρχουν σε μια περιοχή του x και είναι συνεχείς στο x, τότε η F είναι παραγωγίσιμη στο x και

$$F'(x) e_j = \sum_{i=1}^{n} \frac{\partial f_i(x)\, e_j}{\partial x_j} \quad 1 \le j \le n$$

Απόδειξη: R.G. Bartle, "The Elements of Real Analysis", 2nd edition John Whiley & sons 1976 Theorem 39.9 (σελ. 355)

Ε3. Υπολογιστικά Μαθηματικά Ι

8/10/2018

Ορισμός 4: Μια παραγωγίσιμη (ή διαφορίσιμη) απεικόνιση $F: D \subset \mathbb{R}^n \longrightarrow \mathbb{R}^m$ ονομάζεται συνεχώς παραγωγίσιμη στο D αν η $F': D \longrightarrow L(\mathbb{R}^n, \mathbb{R}^m)$ είναι συνεχής απεικόνιση.
Για αυτό απαιτείται το εξής: Δοθέντος ενός $x \in D$ και ενός $\varepsilon > 0$ υπάρχει ένα $\delta > 0$ έτσι ώστε $\|F'y - F'x\| < \varepsilon$, αν $y \in D$ και $\|x - y\| < \delta$.
Στη περίπτωση αυτή λέμε ότι η F είναι απεικόνιση C' ή C^1 ή ότι $F \in C'(D)$ ή $F \in C^1(D)$

Θεώρημα 1: Έστω $F: D \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}^m$, τότε $F \in C'(D)$ ή $F \in C^1(D)$ αν και μόνο αν οι μερικές παράγωγοι: $\frac{\partial f_i(x)}{\partial x_j}$, $1 \leq i \leq m$, $1 \leq j \leq n$ υπάρχουν και είναι συνεχείς στο D

Απόδειξη: W. Rudin, "Principles of Mathematical Analysis", 3rd edition, McGraw-Hill, 1976 Theorem 9.21 (σελ: 219-220)

$$\text{Άρα } [F'(x)] = \begin{bmatrix} \frac{\partial f_1}{\partial x_1}(x) & \frac{\partial f_1}{\partial x_2}(x) & \cdots & \frac{\partial f_1}{\partial x_n}(x) \\ \frac{\partial f_2}{\partial x_1}(x) & \frac{\partial f_2}{\partial x_2}(x) & \cdots & \frac{\partial f_2}{\partial x_2}(x) \\ \vdots & \vdots & & \\ \frac{\partial f_m}{\partial x_1}(x) & \frac{\partial f_m}{\partial x_2}(x) & \cdots & \frac{\partial f_m}{\partial x_n}(x) \end{bmatrix} = J(x)$$

(Ιακωβιανός πίνακας της F)

Μάθημα 4ο   Ε3. Υπολογιστικά Μαθηματικά Ι   10/10/2018

Παραδείγματα

1. Έστω $F:\mathbb{R}^n \to \mathbb{R}^m$. Η F είναι παραγωγίσιμη σε ένα $x\in\mathbb{R}^n$ αν υπάρχει ένας γραμμικός τελεστής $A_x:\mathbb{R}^n \to \mathbb{R}^m$ τέτοιος ώστε

$$\lim_{h\to 0}\frac{\|F(x+h)-F(x)-A_x h\|}{\|h\|}=0$$

($h\in\mathbb{R}^n$, όπου στον αριθμητή η $\|\ \|$ είναι του $\mathbb{R}^m$, ενώ στον παρανομαστή του $\mathbb{R}^n$)

Ο A_x παριστάνεται από τον $m\times n$ Ιακωβιανό πίνακα της

$F(x)=(f_1(x),f_2(x),\dots,f_m(x))^T$ $J(x)=\left[\frac{\partial f_i}{\partial x_j}(x)\right]_{\substack{1\le j\le n\\ 1\le i\le m}}$

2. Έστω $f:\mathbb{R}^n\to\mathbb{R}^1$ συναρτησοειδές (ειδική περίπτωση του προηγούμενου με $m=1$). Τότε για $x\in\mathbb{R}^n$, η $f'(x)\in L(\mathbb{R}^n,\mathbb{R}^1)$ υπάρχει αν ισχύει

$$\lim_{h\to 0}\frac{|f(x+h)-f(x)-f'(x)h|}{\|h\|}=0 \qquad (h\in\mathbb{R}^n)$$

Η $f'(x)$ παριστάνεται από τον $1\times n$ πίνακα (πίνακας γραμμή)

$$[f'(x)]=\left[\frac{\partial f(x)}{\partial x_1},\frac{\partial f(x)}{\partial x_2},\dots,\frac{\partial f(x)}{\partial x_n}\right]=[\nabla f(x)]^T$$

3. Έστω $B\in L(\mathbb{R}^n,\mathbb{R}^m)$, $b\in\mathbb{R}^m$ και ορίζουμε $F:\mathbb{R}^n\to\mathbb{R}^m$ $F(x)=Bx+b$

Για $x\in\mathbb{R}^n$, ισχύει $F'(x)=B$;

Έχουμε $\lim_{h\to 0}\frac{\|F(x+h)-F(x)-F'(x)h\|}{\|h\|}=\lim_{h\to 0}\frac{\|B(x+h)+b-Bx-b-Bh\|}{\|h\|}=0 \Rightarrow F'(x)=B$

4. Έστω $\varphi:\mathbb{R}^n\to\mathbb{R}^1$, $\varphi(x)=\frac{1}{2}(Ax,x)-(b,x)$

$A\in\mathbb{R}^{n\times n}$, $b\in\mathbb{R}^n$ και $(\cdot,\cdot)$ το σύνηθες εσωτερικό γινόμενο

Στο παράδειγμα 2, είδαμε ότι $[\varphi'(x)]=[\nabla\varphi(x)]^T$. Ποιος είναι ο τύπος για $\varphi'(x)$;

Έστω $y\in\mathbb{R}^n$ $y\neq 0$ θέτουμε $h=t\cdot y$ (οπότε $h\to 0 \Leftrightarrow t\to 0$) και έχουμε

$$\lim_{t\to 0}\frac{|\varphi(x+ty)-\varphi(x)-\varphi'(x)(ty)|}{\|ty\|}=\lim_{t\to 0}\frac{|\varphi(x+ty)-\varphi(x)+t\varphi'(x)y|}{|t|\,\|y\|}=$$

$$=\frac{1}{\|y\|}\lim_{t\to 0}\left|\frac{\varphi(x+ty)-\varphi(x)}{t}-\varphi'(x)y\right|=0$$

Τώρα $\varphi(x+ty)-\varphi(x)=\frac{1}{2}(A(x+ty),x+ty)-(b,x+ty)-\frac{1}{2}(Ax,x)+(b,x)$
$=\frac{1}{2}(Ax,x)+\frac{1}{2}t(Ax,y)+\frac{1}{2}t(Ay,x)+\frac{1}{2}t^2(Ay,y)-(b,x)-t(b,y)-\frac{1}{2}A(x,x)+(b,x)$

$\Rightarrow \dfrac{\varphi(x+ty)-\varphi(x)}{t}=\frac{1}{2}(Ax,y)+\frac{1}{2}(Ay,x)+\frac{1}{2}t(Ay,y)-(b,y)$

Θέτουμε $\varphi'(x)y=\frac{1}{2}(Ax,y)+\frac{1}{2}(Ay,x)-(b,y)=\frac{1}{2}(Ax,y)+\frac{1}{2}(A^Tx,y)-(b,y)$
$\Rightarrow \varphi'(x)y=(\frac{1}{2}(A+A^T)x,y)-(b,y)$
οπότε

$$\frac{1}{\|y\|}\lim_{t\to 0}\left|\frac{\varphi(x+ty)-\varphi(x)}{t}-\varphi'(x)y\right|=\frac{1}{\|y\|}\lim_{t\to 0}\left|\frac{1}{2}t(Ay,y)\right|=0$$

Ελέγχουμε τώρα αν ικανοποιείται ο ορισμός
$\varphi(x+h)-\varphi(x)-\varphi'(x)h=\frac{1}{2}(A(x+h),x+h)-(b,x+h)-\frac{1}{2}(Ax,x)+(b,x)-\frac{1}{2}(Ax,h)-\frac{1}{2}(Ax,h)+(b,h)$

$=\frac{1}{2}(Ax,x)+\frac{1}{2}(Ax,h)+\frac{1}{2}(Ah,x)+\frac{1}{2}(Ah,h)-(b,x)-(b,h)-\frac{1}{2}(Ax,x)+(b,x)-\frac{1}{2}(Ax,h)-\frac{1}{2}(Ah,x)+(b,h)$

$=\frac{1}{2}(Ah,h)$ Άρα $\displaystyle\lim_{h\to 0}\frac{|\varphi(x+h)-\varphi(x)-\varphi'(x)h|}{\|h\|}=\lim_{h\to 0}\frac{|\frac{1}{2}(Ah,h)|}{\|h\|}\le\frac{1}{2}\lim_{h\to 0}\frac{\|Ah\|\cdot\|h\|}{\|h\|}$

$=\displaystyle\frac{1}{2}\lim_{h\to 0}\|Ah\|\le\frac{1}{2}\lim_{h\to 0}\|A\|\,\|h\|=0$

Η ύπαρξη μόνο των μερικών παραγώγων $\frac{\partial f_i}{\partial x_j}(x)$ $1\le i,j\le n$ δεν εγγυάται ότι η F είναι παραγωγίσιμη στο x. Αυτό συνάγεται από το ακόλουθο σημαντικό αποτέλεσμα

Πρόταση 3: Έστω ότι η $F:D\subset\mathbb{R}^n\to\mathbb{R}^n$ είναι παραγωγίσιμη στο σημείο $x\in \mathrm{Int}D$. Τότε η F είναι συνεχής στο x.

Απόδειξη:

Εφόσον $x\in \mathrm{Int}D$ $\exists\ \delta_0>0: \|h\|<\delta_0 \Rightarrow x+h\in D$
Η F είναι παραγωγίσιμη στο x, συνεπώς, για δεδομένο $\varepsilon>0$ $\exists\ \delta>0$, το οποίο μπορούμε να πάρουμε $\le\delta_0$: Για $\|h\|<\delta$ ισχύει
$\dfrac{\|F(x+h)-F(x)-F'(x)h\|}{\|h\|}<\varepsilon \Rightarrow \|F(x+h)-F(x)-F'(x)h\|<\varepsilon\|h\| \Rightarrow$

Ε3. Υπολογιστικά Μαθηματικά Ι

10/10/2018

$$\Longrightarrow \|F(x+h)-F(x)\| - \|F'(x)\|\,\|h\| \leq \|F(x+h)-F(x)\| - \|F'(x)h\|$$
$$\leq \|F(x+h)-F(x)-F'(x)h\| \leq \varepsilon\|h\|$$
$$\Longrightarrow \|F(x+h)-F(x)\| \leq \|F'(x)\|\,\|h\| + \varepsilon\|h\| = \overbrace{(\|F'(x)\|+\varepsilon)}^{c}\,\|h\| = c\|h\|$$

Σταθεροποιώντας το ε, συμπεραίνουμε ότι για δεδομένο $x \in \operatorname{Int} D$ $\exists\, \delta>0$ και $c>0$:
$\|h\|<\delta \Longrightarrow \|F(x+h)-F(x)\| \leq c\|h\|$. Που είναι μια τοπική συνθήκη Lipschitz, δηλαδή ισχυρότερη από την συνέχεια στο x.

Παράδειγμα: Η συνάρτηση $f: \mathbb{R}^2 \longrightarrow \mathbb{R}^1$ που δίνεται από:

$$f(x_1,x_2) = \begin{cases} 0 &, \text{αν } (x_1,x_2)=(0,0) \\ \dfrac{x_1x_2}{x_1^2+x_2^2} &, \text{αν } (x_1,x_2)\neq(0,0) \end{cases}$$

Έχει μερικές παραγώγους $\frac{\partial f}{\partial x_1}$ και $\frac{\partial f}{\partial x_2}$ σε κάθε σημείο του $\mathbb{R}^2$, αλλά η f δεν είναι συνεχής στο $(0,0)$

Πράγματι $$\frac{\partial f}{\partial x_1}(x_1,x_2) = \begin{cases} 0 &, \text{αν } (x_1,x_2)=(0,0) \\ \dfrac{x_2^3-x_1^2x_2}{(x_1^2+x_2^2)^2} &, \text{αν } (x_1,x_2)\neq(0,0) \end{cases}, \quad \frac{\partial f}{\partial x_2}(x_1,x_2) = \begin{cases} 0 &, \text{αν } (x_1,x_2)=(0,0) \\ \dfrac{x_1^3-x_2^2x_1}{(x_1^2+x_2^2)^2} &, \text{αν } (x_1,x_2)\neq(0,0) \end{cases}$$

Επιπλέον αν $x_1=x_2=h$, τότε $\displaystyle\lim_{(x_1,x_2)\to(0,0)} \frac{x_1x_2}{x_1^2+x_2^2} = \lim_{h\to 0}\frac{h^2}{2h^2} = \lim_{h\to 0}\frac{1}{2} = \frac{1}{2}$

Ενώ αν $x_1=-h, x_2=h$, τότε $\displaystyle\lim_{(x_1,x_2)\to(0,0)} \frac{x_1x_2}{x_1^2+x_2^2} = \lim_{h\to 0}\frac{h^2}{-2h^2} = \lim_{h\to 0} -\frac{1}{2} = -\frac{1}{2}$

1.2. Γενική επαναληπτική μέθοδος. Τοπικά Θεωρήματα σύγκλισης. Το Θεώρημα της Συστολής

Θα ασχοληθούμε με την κατασκευή επαναληπτικών μεθόδων, για την προσέγγιση των λύσεων του μη Γραμμικού συστήματος

$$F(x)=0 \qquad (2)$$

όπου $F: D\subset\mathbb{R}^n \longrightarrow \mathbb{R}^n$ είναι μια δεδομένη απεικόνιση. Ξεκινώντας λοιπόν με μια αρχική προσέγγιση $x^0\in\mathbb{R}^n$, οι μέθοδοι μας παράγουν μια ακολουθία $\{x^k\}, k\geq 1$ διανυσμάτων του $\mathbb{R}^n$, τέτοια ώστε $x^k \longrightarrow x^*$ καθώς $k\longrightarrow\infty$, όπου x^* είναι μια λύση του (2)

Μάθημα 5ο Ε3. Υπολογιστικά Μαθηματικά Ι 15/10/2018

Βασικά ερωτήματα που θα μας απασχολήσουν

1. Οι προσεγγίσεις $\{x^k\}, k \geq 1$ θα πρέπει να είναι καλώς ορισμένες
2. Η ακολουθία $\{x^k\}$ $k \geq 1$ συγκλίνει σε μια λύση x^* του (2). Διακρίνουμε θεωρήματα σύγκλισης διάφορων τύπων:

α. Θεωρήματα τοπικής σύγκλισης: Η αρχική προσέγγιση x^0 επιλέγεται σε μια περιοχή S του x^*

β. Θεωρήματα περιορισμένης σύγκλισης: Η αρχική προσέγγιση x^0 επιλέγεται σε ένα, γενικά περιορισμένο σύνολο τιμών

γ. Θεωρήματα ολικής σύγκλισης: Το x^0 επιλέγεται είτε στο $\mathbb{R}^n$ ή τουλάχιστον σε ένα μεγάλο υποσύνολό του.

3. Η ταχύτητα σύγκλισης της ακολουθίας $\{x^k\}, k \geq 1$, στο x^*.
Οι εκτιμήσεις του σφάλματος $\|x^k - x^*\|$ είναι κάποιες φορές πεσιμιστικές, οπότε μας ενδιαφέρει η συμπεριφορά του σφάλματος για μεγάλο k.
4. Πρακτικά θέματα, όπως η επιλογή των αρχικών τιμών, κριτήρια τερματισμού και εκτέλεση των βημάτων του αλγορίθμου.

Κατ' αρχήν θα δούμε τη γενική επαναληπτική μέθοδο.
Υποθέτουμε ένα $x \in \mathbb{R}^n$ είναι λύση του (2) αν και μόνο αν ικανοποιεί την εξίσωση:

$$x = G(x) \qquad (3)$$

όπου $G: D \subset \mathbb{R}^n \to \mathbb{R}^n$ είναι μια άλλη απεικόνιση.
Δηλαδή κάθε λύση της (2) είναι σταθερό σημείο της G
Από την (3) απορρέει άμεσα το επαναληπτικό σχήμα

$$\begin{aligned} &x^0 \in \mathbb{R}^n \text{ δεδομένο} \\ &x^{k+1} = G(x^k), \quad k = 0,1,2,\dots \end{aligned} \qquad (4)$$

Πράγματι, αν $x^k \in D$, $k \geq 0$ $x^k \to x^*$ καθώς $k \to \infty$, $x^* \in D$ και η G είναι συνεχής στο x^*, τότε

$$x^* = \lim_{k\to\infty} x^{k+1} = \lim_{k\to\infty} G(x^k) = G\left(\lim_{k\to\infty} x^k\right) = G(x^*)$$

Ορισμός: Λέμε ότι ένα σημείο $x^* \in D$ είναι σημείο έλξεως για την ακολουθία $\{x^k\}_{k=0}^{\infty}$, που παράγεται από την αναδρομική σχέση $x^{k+1} = G(x^k)$, $k=0,1,2,\dots$ x^0 δεδομένο, αν και μόνο αν υπάρχει ανοιχτή σφαίρα $S(x^*;\delta) \subset D$ (για κάποια νόρμα) κέντρου x^* και ακτίνας δ, τέτοια ώστε αν $x^0 \in S(x^*;\delta)$ τότε
$x^k \in S(x^*,\delta)$ και $x^k \to x^*$ καθώς $k \to \infty$.

Ξεκινάμε με ένα γενικό αποτέλεσμα τοπικής σύγκλισης.

Λήμμα: Έστω $G: D \subset \mathbb{R}^n \to \mathbb{R}^n$ και $x^* \in D$. Υποθέτουμε ότι υπάρχει ανοιχτή σφαίρα $S(x^*,\delta) \subset D$ και σταθερά $\alpha < 1$, τέτοια ώστε:
$\|G(x) - x^*\| \le \alpha \|x - x^*\| \quad \forall\ x \in S(x^*,\delta)$
Τότε το x^* είναι σημείο έλξεως της $\{x^k\}_{k=0}^{\infty}$

Απόδειξη:
Με επαγωγή ως προς k
Έστω $x^0 \in S(x^*,\delta)$. Για $k=1$ έχουμε:
$\|x^1 - x^*\| = \|G(x^0) - x^*\| \le \alpha \|x^0 - x^*\| \overset{\alpha<1}{<} \|x^0 - x^*\| \le \delta$, οπότε $x_1 \in S(x^*;\delta)$
Αν υποθέσουμε ότι για κάποιο k, $x^k \in S(x^*,\delta)$ τότε:
$\|x^k - x^*\| = \|G(x^{k-1}) - x^*\| \le \alpha \|x^{k-1} - x^*\| = \alpha \|G(x^{k-2}) - x^*\| \le \alpha^2 \|x^{k-2} - x^*\| \le \dots \le \alpha^k \cdot \|x^0 - x^*\|$
Οπότε
$\|x^{k+1} - x^*\| = \|G(x^k) - x^*\| \le \alpha \cdot \|x^k - x^*\| \le \alpha^{k+1} \cdot \|x^0 - x^*\| < 1 \cdot \delta = \delta$
άρα $x^{k+1} \in S(x^*;\delta)$. Επιπλέον
$\|x^k - x^*\| \le \alpha^k \|x^0 - x^*\| \xrightarrow{\alpha<1} 0 \cdot \|x^0 - x^*\| = 0 \implies x^k \to x^*$ καθώς $k \to \infty$,
δηλαδή το x^* είναι σημείο έλξεως της $\{x^k\}_{k=0}^{\infty}$.

Το επόμενο αποτέλεσμα οφείλεται στον Ostrowski

Θεώρημα: Υποθέτουμε ότι η $G: D \subset \mathbb{R}^n$ έχει ένα σταθερό σημείο $x^* \in \mathrm{Int}\, D$ και η G έχει παράγωγο $G'(x^*)$ στο x^*. Αν $\rho(G'(x^*)) \le \sigma < 1$, τότε το x^* είναι σημείο έλξεως της ακολουθίας $\{x^k\}_{k=0}^{\infty}$

Απόδειξη:
Θα χρησιμοποιήσουμε το προηγούμενο λήμμα

Ε3. Υπολογιστικά Μαθηματικά Ι 15/10/2018

Από γνωστό λήμμα έπεται ότι, αν $\varepsilon > 0$, τότε υπάρχει $\|\cdot\|$ τέτοια ώστε
$\|G'(x^*)\| \leq \rho(G'(x^*)) + \varepsilon \leq \sigma + \varepsilon$
Επίσης, εφόσον η G είναι παραγωγίσιμη στο x^*, για το ε αυτό, $\exists\ \delta > 0$ τέτοιο ώστε $S(x^*, \delta) \subset D$ και για $x \in S(x^*, \delta)$ ισχύει
$\|G(x) - G(x^*) - G'(x^*)(x - x^*)\| \leq \varepsilon \cdot \|x - x^*\|$
Από τις δύο προηγούμενες ανισότητες, και δεδομένου ότι $x^* = G(x^*)$ έχουμε

$\|G(x) - x^*\| = \|G(x) - G(x^*) - G'(x^*)(x - x^*) + G'(x^*)(x - x^*)\|$
$\leq \|G(x) - G(x^*) - G'(x^*)(x - x^*)\| + \|G'(x^*)(x - x^*)\| \leq \varepsilon \|x - x^*\| + \|G'(x^*)\| \|x - x^*\|$
$\leq \varepsilon \cdot \|x - x^*\| + (\sigma + \varepsilon)\|x - x^*\| = (\sigma + 2 \cdot \varepsilon)\|x - x^*\|$
Αν τώρα επιλέξουμε $0 < \varepsilon < \frac{1-\sigma}{2}$, τότε $\sigma = \sigma + 2\varepsilon < 1$, οπότε
$\|G(x) - x^*\| \leq \alpha \|x - x^*\| \quad \forall\ x \in S(x^*, \delta)$ συνεπώς ικανοποιούνται οι προϋποθέσεις του προηγούμενου λήμματος, άρα το x^* είναι σημείο έλξεως της $\{x^k\}_{k=0}^{\infty}$

Μάθημα 6ο   Ε3. Υπολογιστικά Μαθηματικά Ι   17/10/2018

Το προηγούμενο αποτέλεσμα είναι ένα τυπικό και ενδιαφέρον αποτέλεσμα τοπικής σύγκλισης. Το επόμενο αποτέλεσμα είναι ένα αποτέλεσμα περιορισμένης σύγκλισης.

Ορισμός: Λέμε ότι η απεικόνιση $G: D \subset \mathbb{R}^n \to \mathbb{R}^n$, είναι συστολή σε ένα σύνολο $D_0 \subset D$ αν υπάρχει νόρμα $\| \ \|$ του $\mathbb{R}^n$ και $\alpha < 1$ τέτοια ώστε:

$$\|G(x) - G(y)\| \leq \alpha \|x - y\| \quad \forall x, y \in D_0$$

Θεώρημα (Συστολής): Υποθέτουμε ότι η απεικόνιση $G: D \subset \mathbb{R}^n \to \mathbb{R}$ είναι συστολή, ως προς τη νόρμα $\|\cdot\|$ του $\mathbb{R}^n$ με σταθερά $\alpha < 1$, στο κλειστό σύνολο $D_0 \subset D$ και ότι $G(D_0) \subset D_0$. Τότε ισχύουν τα ακόλουθα:

(i) Η G έχει ένα μοναδικό σταθερό σημείο $x^* \in D_0$

(ii) Για οποιοδήποτε $x^0 \in D_0$, η ακολουθία $\{x^k\}_{k=0}^{\infty}$ που παράγεται από την $x^{k+1} = G(x^k)$, $k = 0, 1, 2, \ldots$ συγκλίνει στο x^*

(iii) $\|x^k - x^*\| \leq \frac{\alpha^k}{1-\alpha} \|x^1 - x^0\|$, $k = 1, 2, \ldots$

(iv) $\|x^k - x^*\| \leq \frac{\alpha}{1-\alpha} \|x^k - x^{k-1}\|$, $k = 1, 2, \ldots$

(v) $\|x^k - x^*\| \leq \alpha^k \|x^0 - x^*\|$, $k = 1, 2, \ldots$

Απόδειξη:

Έστω $x^0 \in D_0$. Εφόσον $G(D_0) \subset D_0$, η ακολουθία $\{x^k\}_{k=0}^{\infty}$ που παράγεται από την $x^{k+1} = G(x^k)$, $k = 0, 1, 2, \ldots$, είναι καλά ορισμένη και ανήκει στο D_0. Θα δείξουμε ότι η $\{x^k\}_{k=0}^{\infty}$ είναι ακολουθία Cauchy. Έχουμε:

$$\|x^{k+1} - x^k\| = \|G(x^k) - G(x^{k-1})\| \overset{\text{G συστολή}}{\leq} \alpha \|x^k - x^{k-1}\| = \alpha \|G(x^{k-1}) - G(x^{k-2})\| \overset{\text{G συστολή}}{\leq} \alpha^2 \|x^{k-1} - x^{k-2}\| = \ldots$$

$\leq \alpha^k \|x^1 - x^0\|$, $k = 1, 2,$ οπότε, για οποιοδήποτε ακέραιο $m \geq 1$, παίρνουμε για $k = 1, 2, \ldots$

$$\|x^{k+m} - x^k\| = \|x^{k+m} - x^{k+m-1} + x^{k+m-1} - x^{k+m-2} + \ldots + x^{k+1} - x^k\| \leq$$

$$\leq \|x^{k+m} - x^{k+m-1}\| + \|x^{k+m-1} - x^{k+m-2}\| + \ldots + \|x^{k+1} - x^k\| \leq \alpha^{k+m-1} \|x^1 - x^0\| + \alpha^{k+m-2} \|x^1 - x^0\| + \ldots + \alpha^k \|x^1 - x^0\|$$

$$= \alpha^k (\alpha^{m-1} + \alpha^{m-2} + \ldots + \alpha^1) \|x^1 - x^0\| = \alpha^k \cdot \frac{1-\alpha^m}{1-\alpha} \|x^1 - x^0\| \leq \alpha^k \frac{1}{1-\alpha} \|x^1 - x^0\|$$

$$\Leftrightarrow \|x^{k+m} - x^k\| \leq \frac{\alpha^k}{1-\alpha} \|x^1 - x^0\|, \quad k = 1, 2, \ldots \quad (6)$$

Συνεπώς, δεδομένου ότι $\alpha < 1$ η ακολουθία $\{x^k\}_{k=0}^{\infty}$ είναι Cauchy, ως προς τη νόρμα $\|\cdot\|$, στο κλειστό σύνολο D_0

Επομένως, $\exists x^* \in D_0$ τέτοιο ώστε $x^k \to x^*$ καθώς $k \to \infty$

Το x^* είναι σταθερό σημείο της G. Πράγματι,

$\|x^*-G(x^*)\| = \|x^*-x^{k+1}+x^{k+1}-G(x^*)\| \leq \|x^*-x^{k+1}\| + \|x^{k+1}-G(x^*)\| = \|x^*-x^{k+1}\|+\|G(x^k)-G(x^*)\|$

$\leq \|x^*-x^{k+1}\| + \alpha\|x^k-x^*\| \longrightarrow 0$ καθώς $k\longrightarrow\infty$

Το x^* είναι και μοναδικό σταθερό σημείο της G στο D_0, γιατί, αν υπάρχει και ένα άλλο σημείο $x^{**}\in D_0$, $x^{**}\neq x^*$, τέτοιο ώστε $x^{**}=G(x^{**})$, τότε

$\|x^*-x^{**}\| = \|G(x^*)-G(x^{**})\| \leq \alpha\|x^*-x^{**}\| < \|x^*-x^{**}\|$ Άτοπο.

Επίσης, θεωρώντας την $g: D\subset\mathbb{R}^n\to\mathbb{R}^n$, $g(x)=\|x-x^k\|$, τότε, προφανώς η g είναι συνεχής, οπότε

$\|x^k-x^*\| = \|x^*-x^k\| = g(x^*) = g(\lim_{m\to\infty}x^{k+m}) = \lim_{m\to\infty}g(x^{k+m}) = \lim_{m\to\infty}\|x^{k+m}-x^k\| \overset{(6)}{\leq} \frac{\alpha^k}{1-\alpha}\|x^1-x^0\|$, $k=1,2,\ldots$

Ξεκινώντας μια καινούρια ακολουθία προσεγγίσεων με $y^0=x^{k-1}$, έχουμε $y^1=G(y^0)=G(x^{k-1})=x^k$. Εφαρμόζοντας τώρα την προηγούμενη εκτίμηση με $k=1$, έχουμε $\|y^1-x^*\|\leq\frac{\alpha^1}{1-\alpha}\|y^1-y^0\| \Longrightarrow \|x^k-x^*\|\leq\frac{\alpha}{1-\alpha}\|x^k-x^{k-1}\|$, $k=1,2,\ldots$

Τέλος, $\|x^k-x^*\| = \|G(x^{k-1})-G(x^*)\| \leq \alpha\|x^{k-1}-x^*\| = \alpha\|G(x^{k-2})-G(x^*)\| \leq$

$\leq \alpha^2\|x^{k-2}-x^*\| = \ldots \leq \alpha^k\|x^0-x^*\|$, $k=1,2,\ldots$

Παρατηρήσεις

1) Αναφορικά με το προηγούμενο Λήμμα Αν $x,y\in S(x^*,\delta)$, τότε

$\|G(x)-G(y)\| = \|G(x)-x^*+x^*-G(y)\| \leq \|G(x)-x^*\| + \|G(y)-x^*\| \overset{\text{συνθήκη Lipschitz}}{\leq} \alpha\|x-x^*\|+\alpha\|y-x^*\|$

$=\alpha(\|x-x^*\|+\|y-x^*\|)$ (1')

Ενώ, αν η G είναι συστολή στο $S(x^*,\delta)$, με σταθερά $\alpha<1$, τότε για $x,y\in S(x^*,\delta)$

$\|G(x)-G(y)\|\leq\alpha\|x-y\|$ (2')

Δεδομένου ότι $\|x-y\|=\|x-x^*+x^*-y\|\leq\|x-x^*\|+\|y-x^*\|$ η (1') έπεται από τη (2'), αλλά δεν ισχύει το αντίστροφο. Συνεπώς η (2') είναι ισχυρότερη από την (1'), δηλαδή την $\|G(x)-x^*\|\leq\alpha\|x-x^*\|$, $\alpha<1$, και αυτό εξασφαλίζει την περιορισμένη σύγκλιση (2') σε αντίθεση με την τοπική (1')

2) Έχουμε $\|x^k-x^{k-1}\| \leq \alpha^{k-1}\|x^1-x^0\| = \alpha^{k-1}\|x^1-x^*+x^*-x^0\| \leq \alpha^{k-1}(\|x^1-x^*\|+\|x^0-x^*\|)$

$=\alpha^{k-1}(\|G(x^0)-G(x^*)\|+\|x^0-x^*\|) \leq \alpha^{k-1}(\alpha\|x^0-x^*\|+\|x^0-x^*\|) = \alpha^{k-1}(1-\alpha)\|x^0-x^*\|$

$\Longrightarrow \frac{\alpha}{1-\alpha}\|x^k-x^{k-1}\| \leq \frac{\alpha^k}{1-\alpha}\|x^1-x^0\| \leq \frac{1+\alpha}{1-\alpha}\alpha^k\|x^0-x^*\|$

Συνεπώς, το φράγμα (iv) είναι καλύτερο από το φράγμα (iii)
Ενώ και τα δύο είναι μεγαλύτερα του φράγματος (v) το πολύ
κατα $\frac{1+\alpha}{1-\alpha}$. Επίσης, το φράγμα (iii) είναι "a priori" (εκ των προτέρων)
Το φράγμα (iv) "a posteriori" (εκ των υστέρων), ενώ το φράγμα (v)
έχει μικρή πρακτική αξία λόγω του x^* στο δεύτερο μέλος

Μάθημα 7ο   Ε3. Υπολογιστικά Μαθηματικά Ι   24/10/2018

1.3 Μέθοδος του Νεύτωνα: Τοπική σύγκλιση και ταχύτητα σύγκλισης

Αυτό που θέλουμε να κάνουμε είναι να γενικεύσουμε στον $\mathbb{R}^n$, δηλαδή στα μη γραμμικά συστήματα, τη γνωστή στον $\mathbb{R}^1$, δηλαδή στις μη γραμμικές εξισώσεις, μέθοδο των Newton-Raphson για την λύση της $f(x)=0$,

x_0 δεδομένο, $x_{k+1} = x_k - \dfrac{f(x_k)}{f'(x_k)} = x_k - [f'(x_k)]^{-1} f(x_k)$, $k=0,1,2,\ldots$

Η μέθοδος αυτή στον $\mathbb{R}^n$ έχει τη μορφή:

x^0 δεδομένο, $x^{k+1} = x^k - [F'(x^k)]^{-1} F(x^k)$, $k=0,1,2,\ldots$

και προσεγγίζει τη λύση της $F(x)=0$

Κατ' αρχήν θα μελετήσουμε εξειδικεύσεις της γενικής επαναληπτικής μεθόδου στην περίπτωση που η G έχει τη μορφή

$G(x) = x - (A(x))^{-1} \cdot F(x)$, (7)

όπου το x είναι σε ένα κατάλληλο υποσύνολο του $\mathbb{R}^n$ και ο $A(x) \in L(\mathbb{R}^n)$ είναι ένας αντιστρέψιμος γραμμικός τελεστής, δηλαδή επαναλήψεις της μορφής:

x^0 δεδομένο $x^{k+1} = x^k - (A(x))^{-1} \cdot F(x^k)$, $k=0,1,2,\ldots$

και κατόπιν θα θεωρήσουμε τη σημαντική εξειδίκευση της μεθόδου για $A(x) = F'(x)$, η οποία μας δίνει τη μέθοδο του Νεύτωνα.
Ξεκινάμε με ένα πολύ βασικό θεώρημα.

Θεώρημα: Υποθέτουμε αν:

(i) Η απεικόνιση $F: D \subset \mathbb{R}^n \to \mathbb{R}^n$ είναι παραγωγίσιμη σ' ένα σημείο $x^* \in \text{Int}D$, όπου $F(x^*)=0$.

(ii) Αν $S_0 \subset D$ είναι μια ανοιχτή περιοχή του x^*, θεωρούμε την απεικόνιση $S \to L(\mathbb{R}^n)$ με $x \to A(x)$, για την οποία υποθέτουμε ότι είναι συνεχής στο x^*, και επίσης ότι ο $A(x^*)$ είναι αντιστρέψιμος.

Τότε:

(1) $\exists$ κλειστή σφαίρα $\bar{S}(x^*, \delta) \subset S_0$, $\delta > 0$, στην οποία η απεικόνιση

$G(x) = x - (A(x))^{-1}F(x)$, $x \in \bar{S}(x^*; \delta)$, είναι καλά ορισμένη.

(2) Η G είναι παραγωγίσιμη στο x^* και

$G'(x^*) = I - (A(x^*))^{-1} F'(x^*)$

Απόδειξη:

(1) Αρκεί να δείξουμε ότι ο $A(x)$ είναι αντιστρέψιμος για $x \in \bar{S}(x^*; \delta)$

Από τη συνέχεια της $x \longrightarrow A(x)$ στο x^*, για κάποιο $\varepsilon > 0$, $\exists\, \delta = \delta(\varepsilon, x^*) > 0$ τέτοιο ώστε $\bar{S}(x^*, \delta) \subset S_0$ και

$\|A(x) - A(x^*)\| < \varepsilon$, $\forall x \in \bar{S}(x^*; \delta)$ (8)

Θέτουμε $A = A(x^*)$ και $B = A(x) - A(x^*)$

Κατ' αρχήν $\exists (A(x^*))^{-1}$. Έστω $\|(A(x^*))^{-1}\| = \beta$. Τότε λόγω της (8)

$\|(A(x^*))^{-1}\| \, \|A(x) - A(x^*)\| < \beta \cdot \varepsilon \quad \forall x \in \bar{S}(x^*, \delta)$

Έτσι αν επιλέξουμε $0 < \varepsilon < \frac{1}{2\beta}$, τότε $\beta \cdot \varepsilon < \frac{1}{2}$, οπότε

$\|(A(x^*))^{-1}\| \, \|A(x) - A(x^*)\| < \frac{1}{2} \;\; \forall x \in \bar{S}(x^*; \delta)$, άρα $A + B = A(x)$ είναι αντιστρέψιμος

για $x \in \bar{S}(x^*, \delta)$ και $\|(A(x))^{-1}\| = \|[A(x^*) + A(x) - A(x^*)]^{-1}\| \le$

$$\le \frac{\|(A(x^*))^{-1}\|}{1 - \|(A(x^*))^{-1}\| \, \|A(x) - A(x^*)\|} \le \frac{\beta}{1 - \frac{1}{2}} = 2\beta \quad \forall x \in \bar{S}(x^*, \delta) \qquad (9)$$

Συνεπώς, η $G(x)$ είναι καλά ορισμένη για $x \in \bar{S}(x^*; \delta)$

(2) Κατ' αρχήν, εφόσον $F(x^*) = 0$, το x^* είναι σταθερό σημείο της G, δηλαδή

$x^* = G(x^*)$ (10)

Θέλουμε να δείξουμε ότι

$$\lim_{x \to x^*} \frac{\|G(x) - G(x^*) - (I - (A(x^*))^{-1} F'(x^*))(x - x^*)\|}{\|x - x^*\|} = 0$$

(Πρέπει να δείξουμε ότι $\|G(x) - \ldots\| \le \gamma \varepsilon \|x - x^*\|, \gamma > 0$)

Η F είναι παραγωγίσιμη στο x^*, επομένως για $0 < \varepsilon < \frac{1}{2\beta}$,

$\exists\, \delta_1 = \delta_1(\varepsilon, x^*) > 0$ τέτοιο ώστε

$\|F(x) - F(x^*) - F'(x^*)(x - x^*)\| < \varepsilon \|x - x^*\| \;\; \forall x \in \bar{S}(x^*; \delta_1)$ (11)

Παίρνοντας $\delta_2 = \min\{\delta, \delta_1\}$ έχουμε για $x \in \bar{S}(x^*; \delta_2)$,

$\|G(x) - G(x^*) - (I - (A(x^*))^{-1} F'(x^*))(x - x^*)\|$

$\overset{(7)}{=} \|x - (A(x))^{-1}F(x) - x^* - x + x^* + (A(x^*))^{-1} F'(x^*)(x - x^*)\|$

Για να δημιουργήσω την (11)

$\overset{(10)}{\underset{F(x^*)=0}{=}} \|-(A(x))^{-1}F(x) + (A(x))^{-1}F(x^*) + (A(x))^{-1}F'(x^*)(x - x^*) - (A(x))^{-1}F'(x^*)(x - x^*) + (A(x^*))^{-1}F'(x^*)(x - x^*)\|$

$= \|-(A(x))^{-1}[F(x) - F(x^*) - F'(x^*)(x - x^*)] + [-(A(x))^{-1} + (A(x^*))^{-1}] \cdot F'(x^*)(x - x^*)\|$

Ε3. Υπολογιστικά Μαθηματικά Ι 24/10/2018

Θέτε να δημιουργήσω (8),(9)

$= \| -(A(x))^{-1}[F(x)-F(x^*)-F'(x^*)(x-x^*)] + [-(A(x))^{-1} \cdot A(x^*)(A(x^*))^{-1} + (A(x))^{-1} \cdot A(x)(A(x^*))^{-1}] F'(x^*)(x-x^*) \|$

$= \| -(A(x))^{-1}[F(x)-F(x^*)-F'(x^*)(x-x^*)] + (A(x))^{-1}[A(x)-A(x^*)](A(x^*))^{-1} \cdot F'(x^*)(x-x^*) \|$

$\leq \|(A(x))^{-1}\| \cdot \|F(x)-F(x^*)-F'(x^*)(x-x^*)\| + \|(A(x))^{-1}\| \, \|A(x)-A(x^*)\| \cdot \|(A(x^*))^{-1}\| \cdot \|F'(x^*)(x-x^*)\|$

$\overset{(8)(9)}{\underset{(11)}{<}} 2 \cdot \beta \cdot \varepsilon \|x-x^*\| + 2\beta^2 \varepsilon \cdot \|F'(x^*)\| \, \|x-x^*\| = \gamma \cdot \varepsilon \|x-x^*\|$, όπου $\gamma = 2\beta + 2\beta^2 \|F'(x^*)\|$

Συνεπώς η G είναι παραγωγίσιμη στο x^* και

$G'(x^*) = I - (A(x^*))^{-1} \cdot F'(x^*)$

Μάθημα 8ο Ε3. Υπολογιστικά Μαθηματικα Ι 31/10/2018

Άμεσο πόρισμα του προηγούμενου είναι το ακόλουθο αποτέλεσμα:

Πόρισμα: Έστω ότι ισχύουν οι προϋποθέσεις του προηγούμενου θεωρήματος και επιπλέον ότι $\rho(G'(x^*)) = \rho(I-(A(x^*))^{-1}F'(x^*)) = \sigma<1$ (ρ: φασματική ακτίνα) Τότε το x^* είναι σημείο έλξεως της ακολουθίας $\{x^k\}_{k=0}^{\infty}$ που παράγεται από την αναδρομική σχέση $x^{k+1} = G(x^k)$, $k=0,1,2,\dots$ x^0 δεδομένα

Απόδειξη:

Κατ' αρχήν $G(x^*) = x^* - (A(x^*))^{-1}\overbrace{F(x^*)}^{0} = x^*$. Επιπλέον από το προηγούμενο θεώρημα η G είναι παραγωγίσιμη στο x^*, ενώ από την υπόθεσή μας $\rho(G'(x^*)) = \sigma<1$, συνεπώς, από το θεώρημα του Ostrowski, το x^* είναι σημείο έλξεως της $x^{k+1} = G(x^k)$, $k=0,1,2,\dots$, x^0 δεδομένο

Το προηγούμενο πόρισμα δείχνει ότι η ιδανική επιλογή για το $A(x)$, ώστε η $x^{k+1} = G(x^k)$, $k=0,1,\dots$ x^0 δεδομένο να συγκλίνει στο x^* είναι $A(x)=F'(x)$

Θεώρημα: Υποθέτουμε ότι:

(1) Η απεικόνιση $F: D\subset\mathbb{R}^n \to \mathbb{R}^n$ είναι παραγωγίσιμη σε μια ανοιχτή περιοχή $S_0 \subset D$ ενός σημείου x^*, όπου $F(x^*)=0$

(2) Η F' είναι συνεχής στο x^* και ο $F'(x^*)$ είναι αντιστρέψιμος

Τότε το x^* είναι σημείο έλξεως της ακολουθίας που παράγει η μέθοδος του Νεύτωνα $x^{k+1} = x^k - (F'(x^k))^{-1}F(x^k)$ $k=0,1,2,\dots$, x^0 δεδομένο

Απόδειξη:

Εφαρμόζουμε το προηγούμενο θεώρημα με $A(x) = F'(x)$. Εφόσον η $A(x) = F'(x)$ είναι συνεχής στο x^* και ο $A(x^*) = F'(x^*)$ είναι αντιστρέψιμος, υπάρχει κλειστή σφαίρα $\bar{S}(x^*,\delta)\subset S_0$, $\delta>0$, στην οποία η $G(x) = x-(F'(x))^{-1}F(x)$ είναι καλά ορισμένη. Επιπλέον, η G είναι παραγωγίσιμη στο x^* και $G'(x^*) = I-(F'(x^*))^{-1}F'(x^*) = I-I = 0$. Επομένως $\rho(G'(x^*))=0$ (το καλύτερο δυνατό), και σύμφωνα με το προηγούμενο πόρισμα το x^* είναι σημείο έλξεως της ακολουθίας που παράγει η μέθοδος του Νεύτωνα.

(λόγω του $\rho(G'(x^*))=0$ περιμένουμε γρήγορη σύγκλιση)

Πρόταση: Υποθέτουμε ότι η $F: D\subset \mathbb{R}^n \to \mathbb{R}^n$ είναι συνεχώς παραγωγίσιμη σε ένα κυρτό σύνολο $D_0 \subset D$ και ότι επιπλέον υπάρχουν σταθερές $\alpha, p \geq 0$ τέτοιες ώστε

$$\|F'(u)-F'(v)\| \leq \alpha \|u-v\|^p, \quad u,v \in D_0$$

Τότε $\|F(y)-F(x)-F'(x)(x-y)\| \leq \alpha \dfrac{\|x-y\|^{p+1}}{p+1} \quad \forall\ x,y \in D_0$

Απόδειξη: Σημειώσεις Δουγαλή, Παράγραφος 2, Πρόταση 4

Θεώρημα: Έστω ότι ισχύουν οι υποθέσεις (1) και (2) του προηγούμενου θεωρήματος.
Τότε: (1) Για το σημείο έλξεως x^* της ακολουθίας που παράγει η μέθοδος του Νεύτωνα $x^{k+1} = x^k - (F'(x^k))^{-1} F(x^k)$, $k=0,1,2,\dots$ x^0 δεδομένο ισχύει

$$\lim_{k\to\infty} \frac{\|x^{k+1}-x^*\|}{\|x^k-x^*\|} = 0$$

(2) Αν επιπλέον $\exists$ σταθερά C_1 τέτοια ώστε $\|F'(x)-F'(y)\| \leq C_1 \|x-y\|$ (12) για κάθε x,y σε μια ανοιχτή περιοχή του x^*, τότε $\exists$ $k_0 \in \mathbb{N}$ και σταθερά C_2 τέτοια ώστε $\|x^{k+1}-x^*\| \leq C_2 \|x^k - x^*\|^2$ για $k \geq k_0$, δηλαδή η σύγκλιση είναι τετραγωνική για k αρκετά μεγάλο

Απόδειξη:
(1) Στη μέθοδο του Νεύτωνα $G(x) = x - (F'(x))^{-1} F(x)$. Η G είναι ορισμένη σε μια κλειστή σφαίρα $\bar{S}(x^*;\delta)$, $\delta>0$, όπου $F(x^*)=0$, $x^* = G(x^*)$, και όπου η G είναι παραγωγίσιμη με $G'(x^*) = I - (F'(x^*))^{-1} F'(x^*) = 0$
Έτσι αν πάρουμε $x^0 \in \bar{S}(x^*,\delta)$
$x^{k+1} - x^* = G(x^k) - G(x^*) - G'(x^*)(x^k-x^*)$ οπότε

$$\lim_{k\to\infty} \frac{\|x^{k+1}-x^*\|}{\|x^k-x^*\|} = \lim_{k\to\infty} \frac{\|G(x^k)-G(x^*)-G'(x^*)(x^k-x^*)\|}{\|x^k-x^*\|} = 0$$

λόγω της παραγωγισιμότητας της G στο x^* και του ότι $x^k \to x^*$ καθώς $k\to\infty$
(2) Έστω ότι για $k \geq k_0$ τα x^k περιέχονται στην ανοιχτή περιοχή του x^* για την οποία ισχύει η (12)
Για ένα τέτοιο k θεωρούμε ως D_0 το ευθύγραμμο τμήμα $[x^k\ x^*]$.
Το D_0 είναι κυρτό και η (12) ισχύει στο D_0 (πιθανώς με διαφορετική σταθερά, έστω C_1') Έτσι μπορεί να εφαρμοστεί η προηγούμενη πρόταση στο D_0 για $\alpha = C_1'$ και $p=1$, οπότε:

Ε3. Υπολογιστικά Μαθηματικά Ι

31/10/2018

$\|F(x^k)-F(x^*)-F'(x^*)(x^k-x^*)\| \le c\cdot\|x^k-x^*\|^2$ (13), όπου $c=c_1'/2$. Έτσι έχουμε:

$x^{k+1}-x^* = x^k-(F'(x^k))^{-1}F(x^k)-x^* = -(F'(x^k))^{-1}F(x^k)+\underbrace{(F'(x^k))^{-1}F'(x^k)}_{I}(x^k-x^*)$

$= -(F'(x^k))^{-1}[F(x^k)-F'(x^k)(x^k-x^*)]$

$\overset{F(x^*)=0}{=} -(F'(x^k))^{-1}[F(x^k)-F(x^*)-F'(x^k)(x^k-x^*)-F'(x^*)(x^k-x^*)+F'(x^*)(x^k-x^*)]$

$= -(F'(x^k))^{-1}[F(x^k)-F(x^*)-F'(x^*)(x^k-x^*)]+(F'(x^k))^{-1}F'(x^k)(x^k-x^*)-(F'(x^k))^{-1}F'(x^*)(x^k-x^*)$

$= -(F'(x^k))^{-1}[F(x^k)-F(x^*)-F'(x^*)(x^k-x^*)]+(F'(x^k))^{-1}[F'(x^k)-F'(x^*)](x^k-x^*)$

$\Rightarrow \|x^{k+1}-x^*\| \le \|(F'(x^k))^{-1}\|\cdot\|F(x^k)-F(x^*)-F'(x^*)(x^k-x^*)\|+\|(F'(x^k))^{-1}\|\cdot\|F'(x^k)-F'(x^*)\|\cdot\|x^k-x^*\|$

και λόγω των (13) και (12),

$\|x^{k+1}-x^*\| \le \|(F'(x^k))^{-1}\|\cdot c\cdot\|x^k-x^*\|^2+\|(F'(x^k))^{-1}\|\, c_1'\,\|x^k-x^*\|^2 = \|(F'(x^k))^{-1}\|(c+c_1')\|x^k-x^*\|^2$

Δεδομένου ότι η $A(x)=F'(x)$ ικανοποιεί τις συνθήκες του πρώτου θεωρήματος αυτής της παραγράφου, όπως δείξαμε στην (9)

$\|(A(x))^{-1}\| \le 2\beta = 2\cdot\|(A(x^*))^{-1}\|$ οπότε

$\|(F'(x^k))^{-1}\| \le 2\cdot\|(F'(x^*))^{-1}\|$, συνεπώς

$\|x^{k+1}-x^*\| \le 2\|(F'(x^*))^{-1}\|(c+c_1')\cdot\|x^k-x^*\|^2 = c_2\|x^k-x^*\|^2$

όπου $c_2 = \underbrace{2\|(F'(x^*))^{-1}\|(c+c_1')}_{=\,3\cdot c_1'\|(F'(x^*))^{-1}\|}$, $c=\frac{c_1'}{2}$, σταθερά ανεξάρτητη του k.