Επιστημονικοί Υπολογισμοί

Broyden - απόδειξη

Σταμάτιος-Άγγελος Ν. Αλεξανδρόπουλος
e-mail: stalexan@ee.duth.gr

https://www.researchgate.net/profile/Stamatios_Aggelos_Alexandropoulos
https://scholar.google.gr/citations?user=mht7W_YAAAAJ&hl=el
http://cilab.math.upatras.gr

Τμήμα Ηλεκτρολόγων Μηχανικών & Μηχανικών Υπολογιστών
Δημοκρίτειο Πανεπιστήμιο Θράκης
Κιμμέρια 67100, Ξάνθη

6 Απριλίου 2021

ΔΗΜΟΚΡΙΤΕΙΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΘΡΑΚΗΣ

Περιεχόμενα I

Αριθμητική Επίλυση συστημάτων μη-γραμμικών εξισώσεων

Ορισμός

Δοθείσης μιας συνάρτησης $F_n = (f_1, f_2, \dots, f_n) : \mathcal{A}_n \subset \mathbb{R}^n \to \mathbb{R}^n$, αναζητούμε ένα σημείο $x^* \in \mathcal{A}_n$ για το οποίο ισχύει $F_n(x^*) = \mathcal{O}_n$.

Ορισμός

Το σημείο $x^* \in \mathcal{A}_n$ στην παραπάνω περίπτωση καλείται ρίζα ή μηδενικό της συνάρτησης F_n ή λύση της εξίσωσης $F_n(x) = \mathcal{O}_n$.

Προσοχή: Η συνάρτηση F_n θεωρούμε πως είναι συνεχώς παραγωγίσιμη συνάρτηση.

Ισοδύναμα:

$$F_n(x) = \mathcal{O}_n \Leftrightarrow \begin{cases} f_1(x) = 0 \\ f_2(x) = 0 \\ \dots \\ f_n(x) = 0 \end{cases} , \quad x = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n \tag{1}$$

Αριθμητική Επίλυση συστημάτων μη-γραμμικών εξισώσεων

Παράδειγμα

Έστω η συνάρτηση $F_2 : \mathcal{A}_2 \subset \mathbb{R}^2 \to \mathbb{R}^2$, *με τύπο:*

$$F_2(x) = \begin{cases} f_1(x_1, x_2) = 2x_1^2 + 3x_2^2 - 1/2 \\ f_2(x_1, x_2) = 3x_1^3 + x_2^2 + \sqrt{2}/2 \end{cases} .$$

Να βρείτε κάποιο $x^* = (x_1^*, x_2^*) \in \mathcal{A}_2$ *για το οποίο να ισχύει* $F_2(x^*) = \mathcal{O}_2 = (0,0)^\top$.

Προσοχή: Η πολυπλοκότητα του προβλήματος ανεβαίνει όταν, αντί για ένα τέτοιο σημείο, αναζητούμε όλα τα σημεία με αυτή την ιδιότητα.

Η μέθοδος του Broyden

Η μέθοδος του Broyden κατατάσσεται στις μεθόδους τοπικής σύγκλισης και η σπουδαιότητά της έγκειται στο γεγονός ότι επιλύει ένα σύστημα μη-γραμμικών εξισώσεων επιλύοντας επαναληπτικά, κατάλληλα γραμμικά συστήματα.

Το πλεονέκτημα της μεθόδου είναι ότι δεν απαιτεί υπολογισμούς μερικών παραγώγων για τον καθορισμό του Ιακωβιανού πίνακα, όπως αυτή του Newton, αλλά χρησιμοποιεί ένα μητρώο το οποίο αναπροσαρμόζεται κατάλληλα σε κάθε επανάληψη.

Γενικά, αν και η μέθοδος του Broyden υπερέχει της μεθόδου του Newton αναφορικά με το υπολογιστικό κόστος (πλήθος συναρτησιακών υπολογισμών και σχετική ακρίβεια σε κάθε επανάληψη), υστερεί στην ταχύτητα σύγκλισης.

Τέλος, η μέθοδος του Broyden αποτελεί γενίκευση της μεθόδου της τέμνουσας(secant method).

Η μέθοδος του Broyden

Γενικά, με βάση τον τύπο του Taylor θεωρούμε το ακόλουθο πρότυπο:

$$m_k(x) = f(x^k) + a_k(x - x^k) \tag{2}$$

το οποίο για κάθε $a_k \in \mathbb{R}$, πληροί τη σχέση:

$$m_k(x^k) = f(x^k) \tag{3}$$

Παρατηρούμε ότι, αν στην παραπάνω σχέση θέσουμε:

$$a_k = f'(x^k) \tag{4}$$

προκείπτει η μέθοδος των Newton-Raphson. Στην περίπτωση όμως που θέλουμε να αντικαταστίσουμε την παράγωγο $f'(x^k)$ ώστε να σχηματίσουμε τη μέθοδο της τέμνουσας, το πρότυπό μας πρέπει να πληροί τη σχέση:

$$m_k(x^{k-1}) = f(x^{k-1}) \tag{5}$$

Η μέθοδος του Broyden

Έτσι, προκύπτει η σχέση:

$$f(x^{k-1}) = f(x^k) + a_k(x^{k-1} - x^k), \tag{6}$$

η οποία γράφεται στη μορφή:

$$a_k(x^k - x^{k-1}) = f(x^k) - f(x^{k-1}), \tag{7}$$

που ονομάζεται εξίσωση της τέμνουσας και παρέχει την προσέγγιση:

$$a_k = \frac{f(x^k) - f(x^{k-1})}{x^k - x^{k-1}}, \tag{8}$$

Η επόμενη επανάληψη της μεθόδου της τέμνουσας πρέπει να πληροί τη σχέση:

$$m_k(x^{k-1}) = 0, \tag{9}$$

από την οποία, με βάση τη σχέση (2), έχουμε:

$$a_k(x^{k-1} - x^k) + f(x^k) = 0. \tag{10}$$

Απόδειξη της μεθόδου

Από αυτή παίρνουμε τη μέθοδο της τέμνουσας:

$$x^{k+1} = x^k - a_k^{-1} f(x^k) \tag{11}$$

όπου $k = 0, 1, \ldots$ και το a_k δίνεται από την προσέγγιση της σχέσης (2).
Τώρα, στην περίπτωση των συναρτήσεων πολλών μεταβλητών, με βάση το αναπτυγμα του Taylor, το αντίστοιχο πρότυπο θα είναι το εξής:

$$M_k(x) = F_n(x^k) + A_k(x - x^k) \tag{12}$$

όπου για κάθε μητρώο $A_k \in \mathbb{R}^{n \times n}$ πληροί τη σχέση:

$$M_k(x^k) = F_n(x^k). \tag{13}$$

Ακολούθως, αν στη σχέση (12) θέσουμε:

$$A_k = J_{F_n}(x^k), \tag{14}$$

όπου J_{F_n} είναι το Ιακωβιανό μητρώο, παίρνουμε τη μέθοδο του Newton. Στην περίπτωση της γενικευμένης μεθόδου της τέμνουσας, το πρότυπο πρέπει να πληροί τη σχέση:

$$M_k(x^{k-1}) = F_n(x^{x-1}). \tag{15}$$

Απόδειξη της μεθόδου

Έτσι, παίρνουμε τη σχέση:

$$F_n(x^{k-1}) = F_n(x^k) + A_k(x^{k-1} - x^k), \tag{16}$$

η οποία γράφεται:

$$A_k(x^k - x^{k-1}) = F_n(x^k) - F_n(x^{k-1}), \tag{17}$$

που τελικώς μπορεί να γραφεί στην ακόλουθη μορφή, η οποία θα μας απασχολήσει περισσότερο στη συνέχεια:

$$A_{k+1}(x^{k+1} - x^k) = F_n(x^{k+1}) - F_n(x^k). \tag{18}$$

Η συνθήκη που μόλις γράψαμε ονομάζεται συνθήκη quasi-Newton(quasi-Newton condition) ή εξίσωση της τέμνουσας (secant equation).

Παρατήρηση Το σύστημα εξισώσεων της συνθήκης quasi-Newton μας παρέχει n εξισώσεις ώστε να καθοριστούν τα άγνωστα $n \times n$ στοιχεία του μητρώου A_{k+1}. Αυτό όμως έχει ως συνέπεια να μην μπορεί να καθοριστεί μοναδικά το μητρώο A_{k+1}. Έτσι, μια πιθανώς καλή επιλογή για το μητρώο A_{k+1} είναι η ελάχιστη τροποποίηση του μητρώου A_k έτσι ώστε:

$$A_{k+1}t = A_k t, \ (x^{k+1} - x^k)^{\mathrm{T}} t = 0, \ \ t \in \mathbb{R}^n, \tag{19}$$

το οποίο μας οδηγεί στον μοναδικό καθορισμό του μητρώου A_{k+1}. Η επιλογή του μητρώου $A_k k + 1$, έτσι ώστε $(A_{k+1} - A_k)t = 0$ για όλα τα t τα οποία είναι ορθογώνια στο $(x^{k+1} - x^k)$, απαιτεί το μητρώο $(A_{k+1} - A_k)$ να είναι βαθμού ένα, $rank(A_{k+1} - A_k) = 1$, δηλαδή της μορφής:

$$A_{k+1} - A_k = u(x^{k+1} - x^k)^{\mathrm{T}}, u \in \mathbb{R}^n. \tag{20}$$

Απόδειξη της μεθόδου

Ο Charles Broyden παίρνοντας τη γενίκευση της σχέσης (11):

$$x^{k+1} = x^k - A^k F_n(x^k), k = 0, 1, \ldots, \tag{21}$$

υπέδειξε ότι η επιλογή των μητρώων A_k είναι προτιμότερο να γίνεται έτσι ώστε ο βαθμός της διαφοράς δύο διαδοχικών μητρώων, $rank(A_{k+1} - A_k)$ να είναι μικρός. Αυτή η σκέψη οδήγησε στην ανάπτυξη των μεθόδων quasi-Newton(με τις οποίες θα ασχοληθούμε στην αμέσως επόμενη ενότητα), οι οποίες χαρακτηρίζονται από τις παρακάτω τρεις ιδιότητες:

(α) $A_k(x^{k+1} - x^k) + F_n(x^k) = 0,$

(β) $A_{k+1}(x^{k+1} - x^k) = F_n(x^{k+1}) - F_n(x^k), \quad k = 0, 1, \ldots$

(γ) $rank(A_{k+1} - A_k) = m \geqslant 1$

όπου, θέτοντας:

$$s^k = x^{k+1} - x^k, \tag{22}$$

$$y^k = F_n(x^{k+1}) - F_n(x^k) \tag{23}$$

και συνδυάζοντας κατάλληλα τις ιδιότητες (α) και (β) παίρνουμε:

$$(A_{k+1} - A_k)s^k = F_n(x^{k+1}), \tag{24}$$

$$F_n(x^{k+1}) = y^k - A_k s^k \tag{25}$$

Μέχρι τώρα για την ανάπτυξη μεθόδων quasi-Newton έχουν προταθεί μόνο οι τιμές $m = 1$ και $m = 2$ των σχέσεων $(a), (b), (g)$ που είδαμε προηγουμένως. Ο Charles Broyden για την περίπτωση που $m = 1$, έδωσε δύο μεθόδους, την καλή μέθοδο (good method), όπως ονόμασε και την κακή μέθοδο (bad method), η οποία υπολείπεται σε επιδόσεις σε σχέση με την πρώτη.

Σύγκλιση της μεθόδου

Θεώρημα

Έστω $F_n : \mathcal{D} \subset \mathbb{R}^n \to \mathbb{R}^n$ συνεχώς παραγωγίσιμη συνάρτηση σε ένα ανοικτό και κυρτό σύνολο $\mathcal{D}$. Υποθέτουμε ότι υπάρχει $x^\star \in \mathbb{R}^n$ και $r, b > 0$, τέτοια ώστε:
$F_n(x^\star) = J_n, \ \ \mathcal{B}(x^\star, r) \subset \mathcal{D}$, όπου $\mathcal{B}(x^\star, r) = \{x \in \mathbb{R}^n | ||x - x^\star|| < r\}$ είναι ανοιχτή σφαίρα με κέντρο $x^\star$ και ακτίνα r, $\ J_{F_n} \in Lip_\gamma(\mathcal{B}(x^\star, r))$, και
το αντίστροφο μητρώο $J_{F_n}(x\star)^{-1}$ υπάρχει και ισχύει ότι $||J_{F_n}(x^\star)^{-1}|| \leqslant \beta$.
Τότε υπάρχουν $\varepsilon > 0$ και $\delta > 0$ τέτοια ώστε, αν ισχύουν οι σχέσεις:

$$||x^0 - x^\star||_2 \leqslant \varepsilon \text{ και } ||A_0 - J_{F_n}||_2 \leqslant \delta,$$

οπότε η ακολουθία $\{x^k\}$ που παράγεται από τη μέθοδο του Broyden είναι καλά ορισμένη και συγκλίνει $q-$υπεργραμμικά στη λύση $x^\star$

Βιβλιογραφία - Αναφορές

- Γραββάνης Γ., Επιστημονικοί υπολογισμοί: Υπολογιστικές μέθοδοι και μαθηματικό λογισμικό, Παπασωτηρίου, 2013.
- Ακρίβης Γ.Δ., Δουγάλης Β.Α., Εισαγωγή στην αριθμητική ανάλυση, Πανεπιστημιακές εκδόσεις Κρήτης, Ηράκλειο, 1997.
- Βραχάτης Μ.Ν., Αριθμητική ανάλυση: Εισαγωγή, Εκδόσεις Κλειδάριθμος, Αθήνα, 2011.
- Βραχάτης Μ.Ν., Αριθμητική Ανάλυση: Υπερβατικές Εξισώσεις, Εκδόσεις Κλειδάριθμος, Αθήνα, 2012.
- Νικόλαος Μισυρλής. Αριθμητική Ανάλυση: Μια Αλγοριθμική Προσέγγιση. Εκδόσεις: Εκδοτική ΕΚΠΑ 2017. Εύδοξος Αριθμητική Ανάλυση

ΔΗΜΟΚΡΙΤΕΙΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΘΡΑΚΗΣ

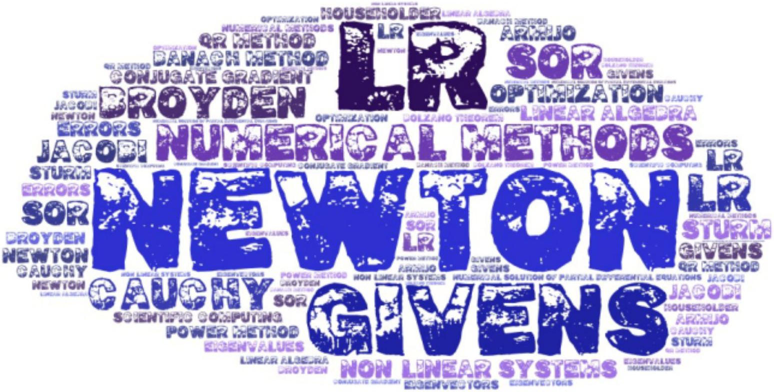
Βιβλιογραφία - Αναφορές

- Gerald, Curtis F, Applied numerical analysis. Pearson Education India, 2004.
- Atkinson, Kendall E, An introduction to numerical analysis. John wiley & sons, 2008.
- Burden, Richard L and Faires, J Douglas, Numerical Analysis, Brooks, Cole, Belmont, CA, 1997.
- Phillips, George M and Taylor, Peter J, Theory and applications of numerical analysis. Elsevier, 1996.
- Davis, Philip J and Rabinowitz, Philip, Methods of numerical integration. Courier Corporation, 2007.

Βιβλιογραφία - Αναφορές

- Herniter, Marc E, Programming in MATLAB. Brooks/Cole Publishing Co., 2000.
- Wolfram, Stephen and others, The MATHEMATICA® book, version 4. Cambridge university press, 1999.
- Gene H. Golub, Charles F. Van Loan, Θεωρία και Υπολογισμοί Μητρώων, Εκδόσεις Πεδίο, 2015.
- Λ. Trefethen, D. Bau, Numerical Linear Algebra, SIAM 1997.
- Β. Δουγαλής, Δ. Νούτσος, Α. Χατζηδήμος, Αριθμητική Γραμμική Άλγεβρα, Πανεπιστημιακές σημειώσεις, Πανεπιστήμιο Ιωαννίνων.

Σας ευχαριστώ πολύ για την προσοχή σας!



ΔΗΜΟΚΡΙΤΕΙΟ
ΠΑΝΕΠΙΣΤΗΜΙΟ
ΘΡΑΚΗΣ