Βελτιστοποίηση

Ιωάννης Γ. Τσούλος

Τμήμα Πληροφορικής και τηλεπικοινωνιών
Πανεπιστήμιο Ιωαννίνων

2022

Περίληψη

Βασικοί ορισμοί

1. **Αντικειμενική συνάρτηση**: $f(x),\ S \subset R^n \rightarrow R$
2. Η συνάρτηση πρέπει να είναι παραγωγίσιμη
3. Βελτιστοποίηση: εύρεση χαμηλότερων τιμών της αντικειμενικής συνάρτησης
4. Θεωρούμε πως εκεί που έχει ελάχιστο η $f(x)$ θα έχει και μέγιστο η $-f(x)$ δηλαδή

$$\max_x \{f(x)\} = -\min_x \{-f(x)\}$$

Όρια συνάρτησης

- Το σύνολο S σπάνια είναι ολόκληρο το R^n
- Πολλές φορές ορίζεται ως

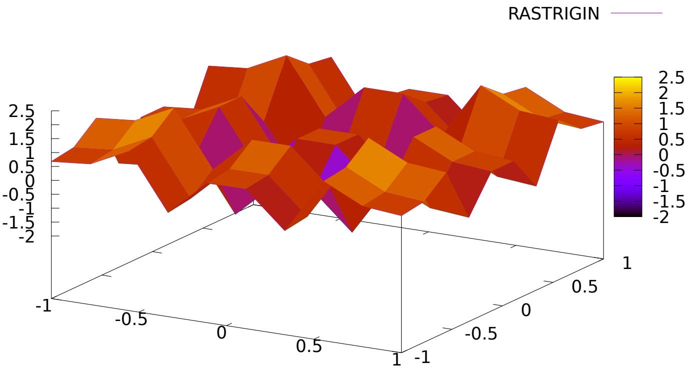
$$S = [a_1, b_1] \otimes [a_2, b_2] \otimes \ldots [a_n, b_n]$$

- Οι τιμές $[a_i, b_i]$ ονομάζονται όρια (bounds) της αντικειμενικής συνάρτησης.
- Παράδειγμα: η συνάρτηση Rastrigin

$$f(x_1, x_2) = x_1^2 + x_2^2 - \cos(18x_1) - \cos(18x_2)$$

with $x \in [-1, 1]^2$

Γράφημα της Rastrigin



Τοπικό ελάχιστο

- Γίνεται αναζήτηση $x^* \in S$, με την ιδιότητα

$$f(x^*) < f(x), \forall x \in S, \|x - x^*\| \leq e$$

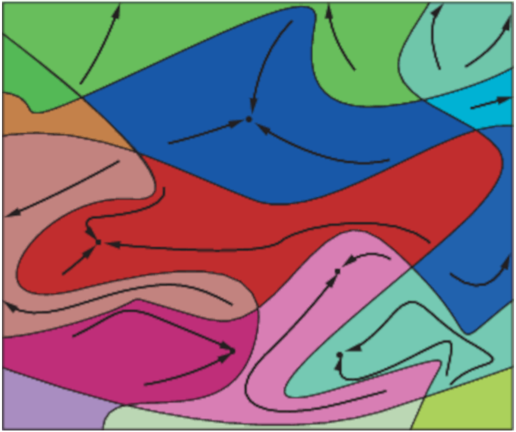
όπου e ένας μικρός θετικός αριθμός.

- Όπως λέμε το x είναι στην **γειτονιά** του x^*

Περιοχή προσέλκυσης

- Αγγλικός όρος Region of Attraction.
- Είναι περιοχές όχι απαραίτητα συνεχόμενες στο πεδίο ορισμού S.
- Έχουν την ιδιότητα πως οποιαδήποτε αναζήτηση από σημεία μιας περιοχής οδηγεί πάντα στο ίδιο τοπικό ελάχιστο.
- Η εύρεση τους είναι σημαντικό στοιχείο στις μεθόδους βελτιστοποίησης, γιατί μειώνεται έτσι ο απαιτούμενος αριθμός αναζητήσεων.

Παράδειγμα περιοχής προσέλκυσης



Ολικό ελάχιστο

1. Στο ολικό ελάχιστο μας ενδιαφέρει η μικρότερη τιμή από όλα τα τοπικά ελάχιστα.
2. Αναζήτηση για x^* με την ιδιότητα

$$f\left(x^*\right) < f(x), \ \forall x \in S$$

3. Σε πολλές περιπτώσεις αναζητούνται και όλα τα τοπικά ελάχιστα μιας συνάρτησης, πχ στις μοριακές διαμορφώσεις ή στις ρίζες εξισώσεων.

Θετικά ορισμένοι πίνακες

1. Θεωρούμε πίνακες συμμετρικούς, $A \in R^{n \times n}$
2. Ένας συμμετρικός πίνακας λέγεται **θετικά ορισμένος** αν: $x^T A x > 0, \ \forall x \neq 0$
3. Ένας συμμετρικός πίνακας λέγεται **θετικά ημιορισμένος** αν: $x^T A x \geq 0, \ \forall x \neq 0$

Αναγκαίες συνθήκες

1. Οι αναγκαίες συνθήκες για να θεωρηθεί ένα σημείο x^* ως ελάχιστο είναι

$$\begin{aligned} \nabla f(x^*) &= 0 \\ \nabla^2 f(x^*) &\geq 0 \end{aligned}$$

2. Στην πρώτη σχέση πρέπει η παράγωγος στο σημείο ενδιαφέροντος να μηδενίζεται.
3. Στην δεύτερη σχέση θα πρέπει ο Εσσιανός Πίνακας να είναι θετικά ημιορισμένος
4. Εσσιανός πίνακας:

$$G_{ij}(x^*) = \frac{\partial^2 f(x^*)}{\partial x_i \partial x_j}$$

Παράδειγμα αναγκαίων συνθηκών

- Η συνάρτηση rastrigin έχει ολικό ελάχιστο στο σημείο $x^* = (0, 0)$
- Παράγωγος είναι:

$$g(x) = (2x_1 + 18\sin(18x_1), 2x_2 + 18\sin(18x_2))$$

- Ο Εσσιανός πίνακας

$$G(x) = \begin{bmatrix} 2 + 324\cos(18x_1) & 0 \\ 0 & 2 + 324\cos(18x_2) \end{bmatrix}$$

- Στο σημείο (0,0) $g(x) = (0, 0)$, $G(x) = \begin{bmatrix} 326 & 0 \\ 0 & 326 \end{bmatrix}$

Ικανές συνθήκες

- Οι ικανές συνθήκες για να θεωρηθεί ένα σημείο x^* ως ελάχιστο είναι

$$\begin{aligned} \nabla f\left(x^*\right) &= 0 \\ \nabla^2 f\left(x^*\right) &> 0 \end{aligned}$$

- Αν ο Εσσιανός πίνακας είναι μονίμως θετικά ορισμένος.
- Συνήθως για λόγους υπολογιστικού κόστους οι μέθοδοι βελτιστοποίησης χρησιμοποιούν μόνο την παράγωγο και ίσως και κάποια αριθμητική προσέγγιση του Εσσιανού πίνακα.

Ανάπτυγμα Taylor

1. Προσέγγιση συναρτήσεων με χρήση παραγώγων.
2. Εφαρμόζεται σε συναρτήσεις που είναι απείρως παραγωγίσιμες. Για παράδειγμα η συνάρτηση $f(x) = e^x$ έχει άπειρες παραγώγους.
3. Για σημεία κοντά σε ένα δοθέν σημείο x_0 ισχύει

$$f(x) = \sum_{n=0}^{\infty} \frac{(x - x_0)^n}{n!} f^{(n)}(x_0)$$

4. Για $x_0 = 0$, η σειρά Taylor ονομάζεται και σειρά Maclaurin.
5. Συνήθως χρησιμοποιείται πεπερασμένος αριθμός όρων για την προσέγγιση της συνάρτησης.

Παραδείγματα σειράς MacLaurin

1. $\cos(x) = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \ldots = \sum_{m=0}^{\infty} \frac{(-1)^m x^{2m}}{(2m)!}$
2. $e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + ... = \sum_{m=0}^{\infty} \frac{x^m}{m!}$

Ορισμοί

1. Η εξίσωση
$$x(\lambda) = x_1 + \lambda\,(x_2 - x_1)$$
δίνει την εξίσωση ευθείας που διέρχεται από τα σημεία x_1, x_2
2. Αν μια ευθεία διέρχεται από το x_1 και είναι παράλληλη σε ένα διάνυσμα s έχει την μορφή
$$x(\lambda) = x_1 + \lambda s$$
3. Η διαδικασία
$$\min_{\lambda} f\,(x + \lambda s)$$
ονομάζεται **γραμμική αναζήτηση**.

Φθίνουσα κατεύθυνση

1. Αναπτύσσοντας την ευθεία κατά Taylor θα έχουμε: $f\left(x_1+\lambda s\right)=f\left(x_1\right)+\lambda s^T \nabla f\left(x_1\right)+\frac{1}{2}\lambda^2 s^T \nabla^2 f\left(x_1\right)s+O\left(\lambda^3\right)$.
2. Η παράγωγος για $\lambda=0$ θα δίνεται από: $\frac{d}{d\lambda}f\left(x_1+\lambda s\right)_{\lambda=0}=s^T \nabla f\left(x_1\right)$
3. Αν η παραπάνω παράγωγος είναι αρνητική, τότε η συνάρτηση στο σημείο x_1 είναι φθίνουσα κατά μήκος του s.
4. Η κατεύθυνση
$$s=-\nabla f\left(x_1\right)$$
είναι φθίνουσα, γιατί: $s^T \nabla f\left(x_1\right)=-\nabla f\left(x_1\right)^T \nabla f\left(x_1\right)<0$ (το εσωτερικό γινόμενο είναι αυστηρά θετικό σε αυτήν την περίπτωση).
5. Οι φθίνουσες κατευθύνσεις παίζουν σημαντικό ρόλο στην βελτιστοποίηση, καθώς έτσι γίνει ευκολότερη η αναζήτηση ελαχίστων.

Γενική μορφή γραμμικής αναζήτησης

1. Δίνεται ένα σημείο $x^{(k)}$
2. Υπολογίζεται μια φθίνουσα κατεύθυνση $s^{(k)}$
3. Ελαχιστοποιείται ως προς λ $(\lambda > 0)$ η συνάρτηση

$$\phi(\lambda) = f\left(x^{(k)} + \lambda s^{(k)}\right)$$

   και εντοπίζεται η βέλτιστη τιμή $\lambda^{(k)}$
4. Τίθεται $x^{(k+1)} = x^{(k)} + \lambda^{(k)} s^{(k)}$

Στο σημείο 3 γίνεται η λεγόμενη **γραμμική αναζήτηση**.

Η μέθοδος της Χρυσής τομής

- Επιλέγονται σημεία x_1, x_2 που να ισαπέχουν από τα άκρα του διαστήματος $[a, b]$.
- Τα βήματα έχουν ως ακολούθως
  1. $\phi = \frac{\sqrt{5}-1}{2}$
  2. **Για** $i = 1, 2, 3, \ldots$ **επαναλαμβάνουμε**
     1. $x_1 = a + (1 - \phi)(b - a)$
     2. $x_2 = a + \phi(b - a)$
     3. **Αν** $f(x_1) < f(x_2)$ **τότε** $b = x_2$
     4. **Αλλιώς** $a = x_1$
     5. **Τέλος - Αν**
  3. **Τέλος - Για**

Παράδειγμα για την f(x)=x(x-3)*(x-3)+2

a	x_1	$f(x_1)$	$f(x_2)$	x_2	b
2.0000	2.7639	2.1540	2.1803	3.2361	**4.0000**
2.0000	2.4721	2.6888	2.1540	2.7639	**3.2361**
2.4721	2.7639	2.1540	2.0091	2.9443	**3.2361**
2.7639	2.9443	2.0091	2.0095	3.0557	**3.2361**
2.8754	2.9443	2.0091	2.0005	2.9868	**3.0557**
2.9443					**3.0557**
- Μετά από 6 επαναλήψεις έχει εντοπιστεί ένα διάστημα [2.9443,3.0557]

Η μέθοδος Fibonacci

- Στην χρυσή τομή το διάστημα αβεβαιότητας μεταβάλλεται με σταθερό ρυθμό σε κάθε επανάληψη.
- Στην μέθοδο Fibonacci καθορίζεται η μείωση από το τρέχον βήμα και από τον αντίστοιχο αριθμό Fibonacci σε αυτό το βήμα.
- Ακολουθία Fibonacci: 1,1,2,3,5,8,13,21,34,55,89,...
- Η ακολουθία Fibonacci αναδρομικά:

$$F_k = F_{k-1} + F_{k-2}$$

  με $F_0 = 1, \;\; F_1 = 1$
- Η ακολουθία δίνεται και σε κλειστή αναλυτική μορφή:

$$F_k = \frac{\left(1+\sqrt{5}\right)^{(k+1)} - \left(1-\sqrt{5}\right)^{(k+1)}}{2^{(k+1)}\sqrt{5}}$$

Παραβολική Παρεμβολή

- Η ιδέα είναι να βρεθεί ένα σήμειο $x_m \in [a, b]$ με $f\left(x_m\right) < \min\left(f(a), f(b)\right)$
- Επίσης να βρεθεί ένα ακόμα σημείο x_s τέτοιο ώστε να ελαχιστοποιείται η μοναδική παραβολή που διέρχεται από τα τρία σημεία: $\left(a, f(a)\right), \left(x_m, f\left(x_m\right)\right), \left(b, f(b)\right)$
- $$x_s = \frac{a+b}{2} - \frac{1}{2}\frac{\frac{f(b)-f(a)}{b-a}}{\frac{\frac{f(x_m)-f(a)}{x_m-a}-\frac{f(b)-f(a)}{b-a}}{x_m-b}}$$

Μέθοδος Armijo

1. Τίθενται $\lambda = 1,\ c_1 \in (0,1),\ \rho \in (0,1)$
2. Αν $f(x + \lambda s) \leq f(x) + c_1 \lambda s^T \nabla f(x)$ τότε
    1. Τερματίζεται η διαδικασία με $\lambda^* = \lambda$
3. Αλλιώς
    1. $\lambda = \rho\lambda$ και μετάβαση στο βήμα 2.

Σύνοψη

- Παρουσιάστηκαν βασικές έννοιες Βελτιστοποίησης.
- Παρουσιάστηκαν έννοιες γραμμικής αναζήτηση.

Βιβλιογραφία I

- Βελτιστοποίηση τεχνικών συστημάτων, Άγγελος Πρωτοπαπάς, 2015, Κάλλιπος, ISBN: 978-960-603-493-0
- Καθολική βελτιστοποίηση: μέθοδοι λογισμικό και εφαρμογές, Ιωάννης Γ. Τσούλος, διδακτορική διατριβή, εθνικό αρχείο διδακτορικών διατριβών.
- Αριθμητικές μέθοδοι βελτιστοποίησης, Ισαάκ Λαγαρής, ιστοσελίδα σημειώσεων διαθέσιμη από http://www.cs.uoi.gr/~lagaris/OPT_UNDER/
- Τεχνικές βελτιστοποίησης, Γεώργιος Α. Ροβιθάκης, 2007, εκδόσεις Τζιόλα, ISBN-13: 978-960-418-141-4 .

Βιβλιογραφία II

Μαθηματική θεωρία βελτιστοποίησης , DingZhu Du, Panos M. Pardalos, Weili Wu, 2005, Εκδόσεις Νέων Τεχνολογιών, ISBN 960-8105-79-X, ISBN-13 978-960-8105-79-9