Βελτιστοποίηση

Ιωάννης Γ. Τσούλος

Τμήμα Πληροφορικής και τηλεπικοινωνιών
Πανεπιστήμιο Ιωαννίνων

2022

Περίληψη

Σχήμα μεθόδων πολλών εκκινήσεων

1. Λήψη N δειγμάτων από την συνάρτηση $f(x)$
2. Επεργασία των δειγμάτων, πχ εκτέλεση μεθόδου ελαχιστοποίησης
3. Έλεγχος για τερματισμό. Αν όχι μετάβαση στο 1
4. Το δείγμα με την χαμηλότερη τιμή είναι και το τελικό αποτέλεσμα.

Δειγματοληψία

1. Το πεδίο ορισμού της συνάρτησης $f(x)$ είναι:

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

2. Η δειγματοληψία είναι βασικό στοιχείο των μεθόδων πολλαπλών εκκινήσεων.
3. Τα δείγματα λαμβάνονται με κάποια τυχαιότητα από το πεδίο ορισμού S

Ομοιόμορφη δειγματοληψία

- Είναι η πλέον συνηθισμένη διαδικασία επιλογής δειγμάτων
- Κάθε συνιστώσα υπολογίζεται από τον τύπο

$$x_i = a_i + (b_i - a_i) \times r$$

όπου r τυχαίος αριθμός στο διάστημα [0,1].

- Όσα περισσότερα δείγματα τόσο καλύτερα καλύπτεται ο χώρος έρευνας.
- **Ωστόσο** πολλά δείγματα ίσως να είναι κοντά και να δώσουν τις ίδιες συναρτησιακές τιμές μετά από ελαχιστοποίηση.

Δειγματοληψία με αποστάσεις

1. Κάθε δείγμα λαμβάνεται ώστε να μην είναι κοντά στα προηγούμενα δείγματα.
2. Χρειάζεται να διατηρείται ιστορικό των δειγμάτων (αυξημένη χρήση μνήμης σε μεγάλα προβλήματα)
3. Το ερώτημα που τίθεται είναι **πόσο** είναι το κοντά.
4. Σε πολλά κριτήρια χρησιμοποιείται ο όγκος του συνόλου S:

$$V(S) = (b_1 - a_1) \times (b_2 - a_2) \times \ldots \times (b_n - a_n)$$

5. Δύο άτομα θεωρούνται κοντά αν απέχουν ένα μικρό κλάσμα του όγκου, πχ

$$\epsilon = \frac{1}{20} V(S)$$

Δειγματοληψία με γείτονες

1. **Θέτουμε** $T_x = \emptyset$, το σύνολο των δειγμάτων από την συνάρτηση
2. **Για** $i = 1, \ldots, N$ **κάνε**
    1. Λήψη $x \in S$
    2. Εύρεση των K κοντινοτέρων γειτόνων για το x, $x_K = \{y_1, y_2, \ldots, y_k\}$ με $y_i \in T_x$
    3. Αν υπάρχει γείτονας με
    $$\|y_i - x\| \leq \epsilon$$
    τότε το x απορρίπτεται και μετάβαση στο 2.a
    4. Αλλιώς $T_x = T_x \cup x$
3. **Τέλος Επανάληψης**

Προβλήματα δειγματοληψίας με γείτονες

1. Η επιλογή της τιμής του K αλλάζει δραματικά την συμπεριφορά της μεθόδου.
2. Απαιτείται εύρεση αποστάσεων και ταξινόμηση
3. Από μόνη της η απόσταση δεν βελτιώνει την επίδοση του αλγορίθμου.
4. Εναλλακτικά μπορεί να υπάρξει ψηφοφορία: ένα δείγμα x θα απορριφθεί αν η πλειοψηφία των γειτόνων του είναι αρκετά κοντά.
5. Εναλλακτικά μπορεί να χρησιμοποιηθεί το επόμενο κριτήριο απόρριψης:
$$\|y_i - x\| \leq \epsilon, \; f(y_i) \leq f(x)$$

Γενικά στοιχεία

- Είναι κρίσιμο να γίνεται τερματισμός της μεθόδου έγκυρα και έγκαιρα
- **Έγκυρα**: να έχει βρεθεί το ολικό ελάχιστο ή όλα τα τοπικά ελάχιστα
- **Έγκαιρα**: να μην αφήνεται η μέθοδος να εκτελείται χωρίς να βρεθεί κάτι καινούριο.

Μέγιστος αριθμός επαναλήψεων

1. Ορίζεται ένας μέγιστος αριθμός επαναλήψεων KMAX.
2. Το κριτήριο είναι $k \geq KMAX$,όπου k η τρέχουσα επανάληψη.
3. Η μέθοδος τερματίζεται ξεπεραστεί αυτός ο αριθμός.
4. Εναλλακτικά ορίζεται ένας μέγιστος αριθμός δειγμάτων NMAX. Ο τερματισμός θα γίνει όταν ξεπεραστεί αυτό το πληθος των δειγμάτων (πχ όταν γίνεται δειγματοληψία ανά ένα).
5. Δεν είναι αποδοτική μέθοδος.

Εύρεση συναρτησιακής τιμής

- Ορίζεται μια τιμή στόχος f^t
- Η μέθοδος ελαχιστοποίησης τερματίζεται όταν $|f^* - f^t| \leq e$
- Είναι χρήσιμη μέθοδος σε ειδικές περιπτώσεις πχ. σε νευρωνικά δίκτυα όπου $f^t = 0$
- Συνήθως συνδυάζεται και μέγιστο αριθμό επαναλήψεων, δηλαδή το κριτήριο θα είναι:

$$k \geq KMAX \text{ OR } |f^* - f^t| \leq e$$

Κριτήριο του Kan

- Διατυπώθηκε το 1987 από τον Kan [6].
- Γίνεται προσπάθεια να εκτιμηθεί το πλήθος των συνολικών τοπικών ελαχίστων.
- Έστω w το πλήθος των διαφορετικών τοπικών ελαχίστων που έχουν βρεθεί και t το πλήθος των συνολικών δειγμάτων ή των τοπικών ελαχιστοποιήσεων.
- Γίνεται εκτίμηση του πλήθους των τοπικών ελαχίστων από την σχέση:

$$\hat{w} = \frac{w(t-1)}{t-w-2}$$

- Η μέθοδος θα τερματιστεί αν

$$\tilde{w} - w < \frac{1}{2}$$

Προβλήματα του κριτηρίου του Kan

- Θα πρέπει να ισχύει $t > 2w^2 + 3w + 2$
- Για να τερματιστεί η μέθοδος θα πρέπει τα δείγματα (τοπικές ελαχιστοποιήσεις) να είναι ανάλογες του τετραγώνου των ελαχίστων.
- Πχ για 100 ελάχιστα θέλουμε 10000 αναζητήσεις.

Κριτήριο DoubleBox

- Σε κάθε επανάληψη k γίνεται καταμέτρηση των συνολικών ελαχίστων w_k
- Υπολογίζεται η διακύμανση της ποσότητας $\sigma\left(w_k\right)$
- Επειδή το πλήθος των ελαχίστων είναι πεπερασμένο, η πόσοτητα αυτή θα μειώνεται.
- Η μέθοδος θα τερματιστεί όταν

  $$\sigma\left(w_k\right) \leq \frac{1}{2}\sigma\left(w_{klast}\right)$$

  όπου klast είναι η τελευταία επανάληψη στην οποία βρέθηκε νέο ελάχιστο.
- Αποδίδει θεαματικά ακόμα και σε πολύπλοκες μεθόδους όπως Genetic Algorithms, PSO κτλ.

Pure Random Search

1. Ορίζεται ένας αριθμός δειγμάτων N
2. Αρχικοποίηση αριθμού επαναλήψεων k=0
3. Ορισμός $f^* = \infty$
4. Για i=1,..,N κάνε
    1. Λήψη ενός νέου δείγματος x_i
    2. Αν $f(x_i) \leq f^*,\ x^* = x_i, f^* = f(x_i)$
5. Τέλος - Επανάληψης
6. k=k+1
7. Έλεγχος κριτηρίου τερματισμού. Αν δεν ισχύει τότε μετάβαση στο βήμα 4.

Προβλήματα Pure Random Search

- Πρέπει να ληφθούν πολλά δείγματα για να καλυφθεί ο χώρος έρευνας
- Τις περισσότερες φορές τα δείγματα x_i δεν είναι τοπικά ελάχιστα.
- Εναλλακτικά ένα σημείο x_i θα ελέγχεται αν $\|g(x_i)\| \leq \epsilon$, όπου g(x) η παράγωγος της συνάρτησης.

MULTISTART

- Είναι παρόμοιας λογικής με την Pure Random Search, αλλά από κάιθε σημείο x_i ξεκινάει μια μέθοδος τοπικής ελαχιστοποίησης $L(x)$
- Θεωρούμε πως η $L(x)$ θα οδηγήσει πάντα σε τοπικό ελάχιστο, άρα δεν χρειάζεται έλεγχος για την παράγωγο.
- Πιθανόν το ίδιο τοπικό ελάχιστο να βρεθεί πολλές φορές.

Multistart (αλγόριθμος)

1. Ορίζεται ένας αριθμός δειγμάτων N
2. Αρχικοποίηση αριθμού επαναλήψεων k=0
3. Ορισμός $f^* = \infty$
4. Για i=1,..,N κάνε
    1. Λήψη ενός νέου δείγματος x_i
    2. $x_i = L\left(x_i\right)$
    3. Αν $f\left(x_i\right) \leq f^*, \ x^* = x_i, f^* = f\left(x_i\right)$
5. Τέλος - Επανάληψης
6. k=k+1
7. Έλεγχος κριτηρίου τερματισμού. Αν δεν ισχύει τότε μετάβαση στο βήμα 4.

MLSL

- Διατυπώθηκε το 1985 από τον Kan [7].
- Από το αρχικό δείγμα σημείων αυτά με την χαμηλότερη τιμή είναι υποψήφια για εκκίνηση της $L(x)$
- Για κάθε υποψήφιο δείγμα βρίσκονται τα κοντινότερα σε αυτό σημεία με χαμηλότερη συναρτησιακή τιμή (Graph Minima) και ανάλογα της απόστασής τους και της συναρτησιακής τους τιμής θα πρέπει να ληφθεί η απόφαση αν θα ξεκινήσει μέθοδος τοπικής ελαχιστοποίησης από αυτά.

Ο αλγόριθμος MLSL

1. Θέσε $k = 0$, ο μετρητής επαναλήψεων.
2. Θέσε $X^* = \emptyset$, είναι το σύνολο με τα τοπικά ελάχιστα.
3. $M_x = \emptyset$.
4. Λήψη Μ δειγμάτων και τοποθέτηση στο M_x.
5. Διατήρηση στο στο M_x ενός ποσοστού $\gamma\%$ από τα αρχικά με την χαμηλότερη συναρτησιακή τιμή.
6. Για κάθε σημείο $x_i \in M_x$ κάνε
    1. Εύρεση του κοντινότερου σημείου $x_j \in M_x$
    2. Αν $\|x_i - x_j\| \leq r_k$ και $f(x_j) \leq f(x_i)$ τότε αγνοούμε το σημείο x_i
    3. Αλλιώς $x_i = L(x_i)\,,\ X^* = X^* \cup x_i$
7. Τέλος Επανάληψης
8. k=k+1
9. Έλεγχος κριτηρίου Τερματισμού. Αν δεν ισχύει μετάβαση στο 3.
10. Επιστροφή του σημείου με χαμηλότερη τιμή στο X^* σαν το συνολικό ελάχιστο.

Η απόσταση r_k στην μέθοδο MLSL

- Η απόσταση με βάση και την επανάληψη k δίνεται από

$$r_k = \pi^{\frac{1}{2}} \left(\sigma \times m(S) \times \Gamma\left(1 + \frac{n}{2}\right) \times \frac{\log(k)}{k} \right)^{\frac{1}{n}}$$

- Το σ είναι σταθερά με συνήθεις τιμές 2,4,...
- Το m(s) είναι

$$m(S) = (b_1 - a_1) \times (b_2 - a_2) \times \ldots \times (b_n - a_n)$$

- Η συνάρτηση $\Gamma(n)$ ορίζεται ως

$$\Gamma(n) = \int_0^\infty t^{n-1} \exp(-t) dt$$

Η μέθοδος Mifinder

- Διατυπώθηκε το 2006 από Tsoulos και Lagaris [8].
- Βασίζεται στην MLSL
- Το κριτήριο απόρριψης βασίζεται στο Κριτήριο της Παραγώγου.
- Η κρίσιμη απόσταση απόρριψης αναπροσαρμόζεται συνεχώς με βάση τις τοπικές αναζητήσεις που έχουν πραγματοποιηθεί.

Το κριτήριο της παραγώγου

- Από Taylor ισχύει για σημεία κοντά σε τοπικό ελάχιστο x^*
$$f(x) \simeq f(x^*) + \frac{1}{2}(x - x^*)^T B^* (x - x^*)$$
B είναι ο Εσσιανός πίνακας.
- Με παραγώγιση παίρνουμε
$$\nabla f(x) \simeq B^* (x - x^*)$$
- Αυτό ισχύει και για y κοντά στο ελάχιστο x^*
$$\nabla f(y) \simeq B^* (y - x^*)$$
- Με αφαίρεση κατά μέλη:
$$(x - y)^T (\nabla f(x) - \nabla f(y)) \simeq (x - y)^T B^* (x - y)$$
- Επειδή ο Εσσιανός είναι θετικά ορισμένος στο σημείο του ελαχίστου έπεται ότι η παραπάνω ποσότητα θα είναι θετικός αριθμός.

Η μέθοδος Minfinder (αλγόριθμος)

1. Θέσε $X^* = \emptyset$, k=0
2. Θέσε $T = \emptyset, \;\; V = \emptyset$
3. Δειγματοληψία N σημείων και προσθήκη στο T.
4. Για κάθε σημειο $x \in T$: αν το x δεν είναι απορριπτέο τότε $V = V \cup x$
5. Αν $\frac{|V|}{N} \leq \frac{1}{2}$ τότε $N = \min\left(N + \frac{N}{10}, NMAX\right)$. Αυτό συμβαίνει για να μην μειωθεί το πλήθος του συνόλου V πάρα πολύ.
6. Για κάθε σημείο $x \in V$ κάνε
    1. Αν x δεν είναι απορριπτέο
        1. $x = LS(x)$
        2. Ενημέρωση κρίσιμης απόστασης r_t
        3. $X^* = X^* \cup x$
    2. Τέλος Αν
7. Τέλος Επανάληψης
8. k=k+1
9. Έλεγχος Κριτηρίου Τερματισμού. Αν δεν ισχύει μετάβαση στο 2.

Η κρίσιμη απόσταση r_t στην Minfinder

- Η κρίσιμη απόσταση r_t δίνεται από:

$$r_t = \frac{1}{M} \sum_{i=1}^{M} \|x_i - L(x_i)\|$$

  όπου M το πλήθος των φορών που έχει εφαρμοστεί η μέθοδος ελαχιστοποίησης $L(x)$

- Δίνει ένα καλό μέτρο της μέσης απόστασης των σημείων από τα τοπικά ελάχιστα της συνάρτησης.

Απορριπτέο σημείο στην Minfinder

- Ένα σημείο θα θεωρηθεί απορριπτέο αν για κάποιο άλλο σημείο $y \in V$ ή $y \in X^*$ ισχύουν ταυτόχρονα
  - $\|x - y\| \leq r_t$
  - $(x - y)^T (\nabla f(x) - \nabla f(y)) > 0$

Σύνοψη

- Μέθοδοι δειγματοληψίας
- Τεχνικές τερματισμού
- Μέθοδοι πολλαπλών εκκινήσεων

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
- Boender, C.G.E., Rinnooy Kan, A.H.G. Bayesian stopping rules for multistart global optimization methods. Mathematical Programming 37, 59–80 (1987)
- Rinnooy Kan, A. H. G., & Timmer, G. T. (1985). The Multi Level Single Linkage Method For Unconstrained And Constrained Global Optimization (No. 2099-2018-3232).
- Ioannis G. Tsoulos, Isaac E. Lagaris, MinFinder: Locating all the local minima of a function, Computer Physics Communications 174, pp. 166-179, 2006.