Βελτιστοποίηση

Ιωάννης Γ. Τσούλος

Τμήμα Πληροφορικής και τηλεπικοινωνιών
Πανεπιστήμιο Ιωαννίνων

2022

Περίληψη

Λήψη τυχαίων σημείων

- Δίνεται η συνάρτηση $f(x), \; x \in S \subset R^n$
- Σε πολλές μεθόδους χρειάζεται να βρεθούν M σημεία από την συνάρτηση, τα οποία να μην σχετίζονται μεταξύ τους.
- Συνήθως λαμβάνουμε τυχαία σημεία από κάποια κατανομή.
- Για κάθε μεταβλητή x_i ισχύει $x \in [a_i, b_i], \; i = 1, \dots, n$
- Πιο συχνά ομοιόμορφη κατανομή:

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

όπου r τυχαίος αριθμός στο [0,1].

Βασικοί ορισμοί

1. Διατυπώθηκε αρχικά από τον Spendley και διορθώθηκε στην συνέχεια από τους Nelder και Mead.
2. Η συνάρτηση f(x) υπολογίζεται στις κορυφές ενός **πολύτοπου** (simplex) με n+1 σημεία, $f_i = f\left(x_i\right), \; i = 0, \ldots, n.$
3. Χρειάζεται **ταξινόμηση** των σημείων: $f_0 < f_1 < f_2 < \ldots < f_n$
4. Ορίζουμε το **κεντροειδές** των n σημείων:

$$x_c = \frac{1}{n} \sum_{i=0}^{n-1} x_i$$

5. Ορίζουμε το **κατοπτρικό** σημείου του x_n ως προς x_c

$$x_r = x_c + a \times \left(x_c - x_n\right)$$

(a συντελεστής ανάκλασης)

Αποφάσεις στην μέθοδο Simplex

1. $f_0 \leq f_r \leq f_{n-1}$
2. $f_r \leq f_0$
3. $f_r > f_{n-1}$

Πρώτη περίπτωση (εισαγωγή)

- Το σημείο f_r γίνεται αποδεκτό στο πολύτοπο και εισάγεται σε αυτό.
- Γίνεται πάλι ταξινόμηση των σημείων:
  $f_0 < f_1 < f_2 < \ldots < f_n$

Δεύτερη περίπτωση (διάταση)

- Εκτελείται η διάταση: $x_e = x_c + \gamma\,(x_r - x_c)$
- γ είναι ο συντελεστής διάτασης, συνήθως $\gamma = 2$
- Αν $f_e < f_r$, τότε $x_n = x_e$ αλλιώς $x_n = x_r$
- Γίνεται πάλι ταξινόμηση των σημείων: $f_0 < f_1 < f_2 < \ldots < f_n$

Τρίτη περίπτωση (συστολή)

- $$x_k = \begin{cases} x_c + \beta\left(x_r - x_c\right) & f_r < f_n \\ x_c + \beta\left(x_n - x_c\right) & f_r \geq f_n \end{cases}$$
- β είναι ο συντελεστής συστολής (συνήθως $\beta = 0.5$)
- Αν $f_k < \min\left(f_r, f_n\right)$, τότε $x_n = x_k$
- Αλλιώς **συστολή** του πολύτοπου ως εξής:

$$x_i = \frac{1}{2}\left(x_o + x_i\right), \quad i = 1, 2, \ldots, n$$

- Γίνεται πάλι ταξινόμηση των σημείων: $f_0 < f_1 < f_2 < \ldots < f_n$

Κριτήριο τερματισμού

- Ορίζεται η μέση συναρτησιακή τιμή

$$m_f = \frac{1}{n+1}\sum_{i=0}^{n} f_i$$

- Ορίζεται η διακύμανση των συναρτησιακών τιμών

$$\sigma_f^2 = \frac{1}{n}\sum_{i=0}^{n}(f_i - m_f)^2$$

- Η μέθοδος τερματίζεται αν $\sigma_f^2 \leq \epsilon$

Ο αλγόριθμος

1. Είσοδος n+1 σημείων, $x_i, \; i = 0, .., n$
2. Υπολογισμός και ταξινόμηση των συναρτησιακών τιμών f_i
3. Αν ισχύει το κριτήριο τερματισμού, τότε τερματισμός και η έξοδος είναι το f_0
4. Υπολογισμός κεντροειδούς x_c και κατοπτρικού x_r
5. Αν ισχύει $f_0 \leq f_r \leq f_{n-1}$, τότε ΕΙΣΑΓΩΓΗ
6. Αλλιώς αν $f_r \leq f_0$, τότε ΔΙΑΤΑΣΗ
7. Αλλιώς ΣΥΣΤΟΛΗ
8. Ταξινόμηση συναρτησιακών τιμών και μετάβαση στο βήμα 3.

Μέθοδος εναλασσόμενων διευθύνσεων

- Γίνεται μονοδιάστατη βελτιστοποίηση ως προς κάθε διάσταση εισόδου
- Εκτελεί επαναλήψεις όσο βελτιώνεται η τιμή της συνάρτησης
- Είναι αργή τεχνική, ειδικά για μεγάλο αριθμό διαστάσεων.
- Παραβλέπει την πιθανή συσχέτιση μεταξύ των μεταβλητών και έτσι δεν είναι αποτελεσματική.

Μέθοδος εναλασσόμενων διευθύνσεων (αλγόριθμος)

1. Δίνεται ένα αρχικό σημείο x
2. **Θέτουμε** $k = 1$
3. **Για** $i = 1, .., n$ **κάνε**
    1. Εύρεση λ^* που ελαχιστοποιεί την $\phi(x) = f(x + \lambda e_i)$
    2. $x = x + \lambda^* e_i$
4. **Τέλος Για**
5. $k = k + 1$
6. **Αν** δεν έχει μειωθεί η τιμή της $f(x)$ **τερματισμός**
7. **Αλλιώς** μετάβαση στο βήμα 3.

Η μέθοδος ROLL

1. Για κάθε διάσταση $i = 1, 2, ..., n$ υπάρχει και μία κατεύθυνση s_i
2. Η μέθοδος προσαρμόζει την κατεύθυνση σε κάθε επανάληψη
3. Στο τέλος κάνει και μια γραμμική αναζήτηση κατά μήκος $s = (s_1, s_2, ..., s_n)$
4. Μέσω της γραμμικής αναζήτησης στο τέλος λαμβάνει υπόψην της και τις πιθανές συσχετίσεις μεταξύ $x_1, x_2, ..., x_n$

Η μέθοδος Roll (αλγόριθμος)

1. Έστω x το αρχικό σημείο
2. **Θέτουμε** k=1
3. **Για** $i = 1, .., n$ **κάνε**
    1. **Θέτουμε** $x_0 = (x_1, x_2, ..., x_n)$
    2. **Θέτουμε** $x_+ = (x_1, x_2, ..., x_i + s_i, ..., x_n)$
    3. **Θέτουμε** $x_- = (x_1, x_2, ..., x_i - s_i, ..., x_n)$
    4. **Αν** $f_+ < f$, $x = x_+, \ \ s_i = as_i, \ \ a > 1$
    5. **Αλλιώς αν** $f_- < f, \ \ x = x_-, \ \ s_i = -as_i, \ \ a > 1$
    6. **Αλλιώς** $s_i = -\frac{1}{2} \frac{f_+ - f_-}{f_+ + f_- - 2f_0} s_i$
4. **Τέλος Για**
5. Γραμμική αναζήτηση κατά μήκος $s = (s_1, s_2, ..., s_n)$
6. **Αν** δεν έχει μειωθεί η τιμή της $f(x)$ **τερματισμός**
7. **Αλλιώς** μετάβαση στο βήμα 3.

Στοχαστική αναζήτηση

1. Θέτουμε k=0
2. Θέτουμε a μια αρχική τιμή με a>0.
3. Επιλέγεται ένα τυχαίο διάνυσμα $s = (s_1, s_2, ..., s_n)$ με $s_i \in [-1, 1]$
4. Αν $f(x + as) < f(x), \; x = x + as$
5. Αλλιώς αν για ένα πλήθος επαναλήψεων δεν έχει υπάρξει βελτίωση μείωση του a.
6. Θέτουμε $k = k + 1$
7. Μετάβαση στο βήμα 2.

Βασικές αρχές

- Αναπτύχθηκε το 1983 από τον Kirkpatrik κυρίως για συνδυαστικά προβλήματα.
- Μιμείται την φυσική διαδικασία της Ανόπτυσης.
- Στην ανόπτυση ζεσταίνεται ένα μέταλλο σε υψηλή θερμοκρασία και εν συνεχεία μειώνεται η θερμοκρασία μέχρι να φτάσει στο 0.
- Στην άνοδο τα μόρια του μετάλλου κινούνται γρήγορα (αναζήτηση λύσεων) αλλά στην συνέχεια όσο αυτό ψύχεται μειώνεται η ταχύτητά τους.
- Ο τρόπος που πέφτει η θερμοκρασία ονομάζεται cooling schedule.

Τεχνικές μείωσης θερμοκρασίας

- Θεωρούμε πως ξεκινάμε πάντα από μια μεγάλη θερμοκρασία T_0
- k ειναι η επανάληψη του αλγορίθμου.
- Εκθετική μείωση:

$$T_k = T_0 a^k, \ 0.8 \leq a \leq 0.9$$

- Logarithmical multiplicative cooling:

$$T_k = \frac{T_0}{1 + a \log(1 + k)}$$

- Linear multiplicative cooling:

$$T_k = \frac{T_0}{1 + ak}$$

- Quadratic multiplicative cooling:

$$T_k = \frac{T_0}{1 + ak^2}$$

Ο αλγόριθμος (μία εκδοχή)

1. **Θέτουμε** $k = 0, \ T_0 > 0$
2. **Έστω** x_0 το αρχικό σημείο.
3. **Έστω** $N_{\text{eps}} > 0$ ένας θετικός ακέραιος.
4. **Έστω** e$>$ 0, ένας μικρός θετικός δεκαδικός αριθμός.
5. **Για** $i = 1, ..., N_{\text{eps}}$ **επανέλαβε**
    1. **Έστω** y ένα νέο δείγμα.
    2. **Αν** $f(y) \leq f\left(x_k\right) \ x_{k+1} = y$
    3. **Διαφορετικά** $x_{k+1} = y$ με πιθανότητα $\min\left\{1, \exp\left(-\frac{f(y)-f(x_k)}{T_k}\right)\right\}$
6. **Τέλος Για**
7. **Ενημέρωση** θερμοκρασίας T_k σύμφωνα με τον μηχανισμό ψύξης.
8. k=k+1
9. **Αν** $T_k \leq e$ **τερματισμός**
10. **Αλλιώς** μετάβαση στο βήμα 5.

Σύνοψη

- Η μέθοδος simplex
- Η μέθοδος roll
- Τυχαία αναζήτηση
- Τεχνικές Simulated Annealing

Βιβλιογραφία I

- KRISHNA B. MISRA, Reliability optimization through sequential simplex search, International Journal of Control Volume 18, 1973.
- Stanley N. Deming and Lloyd R. Parker Jr and M. Bonner Denton, A Review of Simplex Optimization in Analytical Chemistry, C R C Critical Reviews in Analytical Chemistry **7**, pp 187-202, 1978.
- S. KirkpatrickC. D. Gelatt, Jr. and M. P. Vecchi, Optimization by Simulated Annealing, Science **220**, pp. 671-680, 1983.