Μέθοδοι πολυδιάστατης ελαχιστοποίησης χωρίς παραγώγους

Δ. Γ. Παπαγεωργίου
Τμήμα Μηχανικών Επιστήμης Υλικών
Πανεπιστήμιο Ιωαννίνων

dpapageo@cc.uoi.gr
http://pc164.materials.uoi.gr/dpapageo

Μέθοδοι ελαχιστοποίησης πολυδιάστατων συναρτήσεων

Μέθοδοι χωρίς παραγώγους

Ενδεικτικά:

- Μη γραμμική μέθοδος Simplex
- Πολυκατευθυντική αναζήτηση (MDS)
- Αναζήτηση προτύπου

Πλεονεκτήματα:

- Χρησιμοποιούνται μόνο τιμές της συνάρτησης, όχι των παραγώγων.
- Μπορούν να εφαρμοστούν σε ασυνεχείς συναρτήσεις ή συναρτήσεις με θόρυβο.

Μειονεκτήματα:

- Χαμηλή απόδοση όταν έχουμε πολλές μεταβλητές, τυπικά $n > 10$

Μέθοδοι με παραγώγους

Ενδεικτικά:

- Μέθοδος της πιο απότομης καθόδου
- Μέθοδοι συζυγών διευθύνσεων
- Μέθοδος Newton
- Μέθοδος Gauss-Newton
- Μέθοδοι Quasi-Newton

Πλεονεκτήματα:

- Πιο αποδοτικές από μεθόδους χωρίς παραγώγους.
- Υπάρχουν θεωρητικές αποδείξεις σύγκλισης.

Μειονεκτήματα:

- Απαιτούνται οι παράγωγοι της συνάρτησης.
- Δυσκολία εφαρμογής σε μη συνεχείς συναρτήσεις.

Μέθοδος Simplex

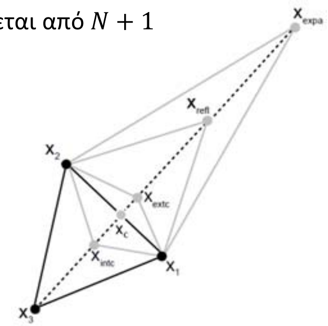
Simplex ονομάζουμε την κυρτή θήκη που σχηματίζεται από $N + 1$ σημεία στον N-διάστατο χώρο.

1. Επιλέγουμε $N + 1$ σημεία.
   $$\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3, \cdots, \boldsymbol{x}_{N+1}$$
2. Υπολογίζουμε τις τιμές της συνάρτησης στα σημεία αυτά και τα αναδιατάσουμε έτσι ώστε:
   $$f(\boldsymbol{x}_1) \le f(\boldsymbol{x}_2) \le \cdots \le f(\boldsymbol{x}_N) \le f(\boldsymbol{x}_{N+1})$$

Καλύτερο σημείο

Δεύτερο χειρότερο σημείο

Χειρότερο σημείο

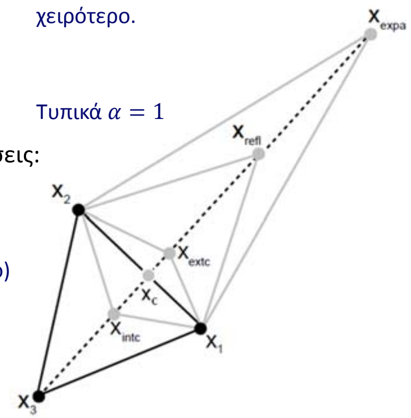


Η βασική ιδέα

Να αντικαταστήσουμε το χειρότερο σημείο με ένα που έχει μικρότερη τιμή συνάρτησης

Μέθοδος Simplex

3. Επανάληψη για $k = 1,2,\cdots$
   a. Υπολογίζουμε το κέντρο του Simplex:
   $$\boldsymbol{x}_c = \frac{1}{N}\sum_{i=1}^{N} \boldsymbol{x}_i$$
   Μέση τιμή όλων των σημείων εκτός από το χειρότερο.
   b. Κατοπτρισμός του χειρότερου σημείο διαμέσου του κέντρου:
   $$\boldsymbol{x}_{refl} = \boldsymbol{x}_c + a(\boldsymbol{x}_c - \boldsymbol{x}_{N+1})$$
   Τυπικά $a = 1$
   c. Εξετάζουμε τις παρακάτω περιπτώσεις:
      I. $f(\boldsymbol{x}_1) \le f(\boldsymbol{x}_{refl}) \le f(\boldsymbol{x}_N)$
      II. $f(\boldsymbol{x}_{refl}) \le f(\boldsymbol{x}_1)$
      (Βρήκαμε ένα νέο καλύτερο σημείο)
      III. $f(\boldsymbol{x}_N) \le f(\boldsymbol{x}_{refl})$
      (Βρήκαμε ένα νέο χειρότερο σημείο)



Μέθοδος Simplex

I. $f(\boldsymbol{x}_1) \leq f(\boldsymbol{x}_{refl}) \leq f(\boldsymbol{x}_N)$ ⇒ Το $\boldsymbol{x}_{refl}$ γίνεται αποδεκτό ως νέο σημείο του simplex και το $\boldsymbol{x}_{N+1}$ απορρίπτεται.

II. $f(\boldsymbol{x}_{refl}) \leq f(\boldsymbol{x}_1)$ ⇒ Ευκαιρία να κινηθούμε σε ακόμη χαμηλότερες τιμές της συνάρτησης.

Εφαρμόζουμε επέκταση:

$$\boldsymbol{x}_{expa} = \boldsymbol{x}_c + \gamma(\boldsymbol{x}_{refl} - \boldsymbol{x}_c)$$

Τυπικά $\gamma = 2$

Αν $f(\boldsymbol{x}_{expa}) \leq f(\boldsymbol{x}_{refl})$ τότε δεχόμαστε το $\boldsymbol{x}_{expa}$ αλλιώς δεχόμαστε το $\boldsymbol{x}_{refl}$

Το $\boldsymbol{x}_{N+1}$ απορρίπτεται.

Μέθοδος Simplex

III. $f(\boldsymbol{x}_N) < f(\boldsymbol{x}_{refl})$ ⇒ Εφαρμόζουμε συστολή

$$\boldsymbol{x}_{contr} = \begin{cases} \boldsymbol{x}_{extc} = \boldsymbol{x}_c + \beta(\boldsymbol{x}_{refl} - \boldsymbol{x}_c) & \text{αν} \quad f(\boldsymbol{x}_N) < f(\boldsymbol{x}_{refl}) < f(\boldsymbol{x}_{N+1}) \\ \boldsymbol{x}_{intc} = \boldsymbol{x}_c + \beta(\boldsymbol{x}_{N+1} - \boldsymbol{x}_c) & \text{αν} \quad f(\boldsymbol{x}_{N+1}) \leq f(\boldsymbol{x}_{refl}) \end{cases}$$

Εξωτερική συστολή

Εσωτερική συστολή

Τυπικά $\beta = 0.5$

Αν $f(\boldsymbol{x}_{contr}) < f(\boldsymbol{x}_{refl})$ και $f(\boldsymbol{x}_{contr}) < f(\boldsymbol{x}_{N+1})$ τότε

κρατάμε το $\boldsymbol{x}_{contr}$

αλλιώς

εφαρμόζουμε σμίκρυνση όλου του simplex

$$\boldsymbol{x}_i = \frac{1}{2}(\boldsymbol{x}_i + \boldsymbol{x}_1) \quad \text{για} \quad i = 2, \cdots, N+1$$

Μέθοδος Simplex

Κριτήρια τερματισμού

$$f(\boldsymbol{x}_{N+1}) - f(\boldsymbol{x}_1) < \varepsilon$$

$$\frac{1}{\mathrm{N}+1}\sum_{i=1}^{N+1}\left|f(\boldsymbol{x}_i) - \bar{f}\right| < \varepsilon$$

όπου

$$\bar{f} = \frac{1}{N+1}\sum_{i=1}^{N+1} f(\boldsymbol{x}_i)$$