

Άσκηση 3: Νευροασαφής Έλεγχος και Εφαρμογές 2021-2022

Μέρος 1: Αλγόριθμοι Βελτιστοποίησης

Στην άσκηση αυτή θα ασχοληθούμε με κάποια υπολογιστικά θέματα στην αριθμητική βελτιστοποίηση.

Θέμα 1. Θεωρήστε τη συνάρτηση

$$f(x_1, x_2) = x_1^2 + (x_2 - 1)^2 + (x_1 - x_2)^4.$$

Εξετάστε τη σύγκλιση της μεθόδου gradient descent για διάφορα μεγέθη βήματος. Συγκρίνετε με τη μέθοδο Newton.

Θέμα 2. Θεωρήστε τη συνάρτηση:

$$f(x_1, x_2) = Ax_1^2 + \frac{1}{A}x_2^2,$$

όπου $A > 0$. Εξετάστε την ταχύτητα σύγκλισης του αλγορίθμου gradient descent για τις διάφορες τιμές του A . Υπολογίστε το Condition number του Hessian πίνακα. Τι παρατηρείτε; Κατόπιν εφαρμόστε τη μέθοδο της ορμής.

Θέμα 3. Ας υποθέσουμε ότι ο Q είναι ένας συμμετρικός θετικά ημι-ορισμένος πίνακας που παράγεται ακολούθως:

$$Q = AA^T,$$

όπου A ένας τυχαίος πίνακας. Εξετάστε το condition number σαν συνάρτηση της διάστασης του πίνακα. Τι παρατηρείτε;

Θέμα 4: Θεωρήστε τη συνάρτηση:

$$f(x_1, x_2) = \max(x_1 + x_2, 0.9x_1 - 1.1x_2 + 1, -0.8x_1 + 1.2x_2 - 1, 2 - 1.1x_1 - 0.9x_2)$$

Εφαρμόστε τη μέθοδο του gradient descent. Παρατηρείτε σύγκλιση; Αν όχι που το αποδίδετε; Πώς αλλάζει η συμπεριφορά όταν αλλάζει το μήκος του βήματος; Δοκιμάστε την ίδια μέθοδο με βήμα $1/k$.

Θέμα 5: Θεωρήστε τη συνάρτηση:

$$f(x_1, x_2) = \frac{1}{4} \sum_{n=1}^4 [(x_1 - a_i)^2 + (x_2 - b_i)^2]$$

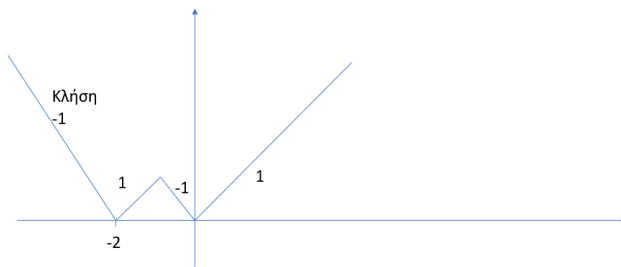
Υποθέστε ότι σε κάθε επανάληψη μπορούμε να υπολογίσουμε μόνο ένα από τα τέσσερα

$$d_i = \frac{1}{2} [(x_1 - a_i) \quad (x_2 - b_i)]^T.$$

Εφαρμόστε τη μέθοδο του stochastic gradient descent.

Μέρος 2: Απλά Νευρωνικά Δίκτυα

Θέμα 1. Κατασκευάστε ένα νευρωνικό δίκτυο με 1 κρυφό στρώμα (με ενεργοποίηση ReLU) και γραμμική έξοδο που να παριστάνει τη συνάρτηση του σχήματος



Θέμα 2. Επεξεργαστείτε το παράδειγμα στον παρακάτω σύνδεσμο:

<https://colab.research.google.com/github/tensorflow/docs/blob/master/site/en/tutorials/keras/classification.ipynb>

Εξετάστε πώς εξαρτώνται τα αποτελέσματα από τη μέθοδο βελτιστοποίησης, τον αριθμό των νευρώνων και των κρυφών στρωμάτων.