Πανεπιστήμιο Ιωαννίνων – Πολυτεχνική Σχολή
Τμήμα Μηχανικών Η/Υ και Πληροφορικής
Μεταπτυχιακό Μάθημα
(Δ3) ΒΕΛΤΙΣΤΟΠΟΙΗΣΗ
Ακαδημαϊκό Έτος 2022-2023

Εργασία 1:

Εύρεση Βέλτιστου Πολυωνύμου

Παπάζης Στέργιος, ΑΜ: 483
Σιουράς Σπύρος, ΑΜ: 485

1 Εισαγωγή

Ο στόχος της εργασίας είναι η εύρεση συντελεστών ενός πολυωνύμου τετάρτου βαθμού που προσεγγίζει μία χρονοσειρά με βέλτιστο τρόπο, δηλαδή να ελαχιστοποιεί το μέσο τετραγωνικό σφάλμα από αυτή πάνω από ένα σύνολο σημείων της με χρήση μεθόδων, αλγορίθμων και τεχνικών από το μάθημα της Βελτιστοποίησης.

Αν ονομάσουμε y την δεδομένη χρονοσειρά και $P_x(t) = \sum_{i=0}^{4} x_i t^i$ για κάθε $x \in \mathbb{R}^5$, η αντικειμενική συνάρτηση της οποίας αναζητούμε το ελάχιστο είναι η:

$$f(x) := \frac{1}{m} \sum_{t=1}^{m} (y_t - P_x(t))^2,$$

όπου $m \in \mathbb{N}$ είναι το πλήθος κάποιων σημείων της χρονοσειράς.

Αυτό το πρόβλημα μπορεί να επιλυθεί αρκετά εύκολα αναλυτικά. Πράγματι, εφόσον η αντικειμενική συνάρτηση είναι ένα πολυώνυμο δευτέρου βαθμού ως προς x, έχει την μορφή

$$f(x) = \frac{1}{2} x^T \cdot M \cdot x + v^T \cdot x + \gamma^2$$

για κάποιες σταθερές $\gamma \in \mathbb{R}$, $v \in \mathbb{R}^5$, $M \in \mathbb{R}^{5\times 5}$ και τα ακρότατά της, αν υπάρχουν, δίνονται ως λύσεις του γραμμικού συστήματος

$$\nabla f(x) = 0 \iff M \cdot x = -v.$$

Αντικαθιστώντας τις τιμές που έχουμε για την χρονοσειρά του προβλήματος και τα πρώτα 25 σημεία βρίσκουμε μοναδικό ελαχιστοποιητή στο σημείο

$$(7.02e^{-01}, -1.61e^{-01}, 2.67e^{-03}, 5.49e^{-04}, -1.80e^{-05}),$$

όπου η συνάρτηση λαμβάνει περίπου την τιμή 0.0122. Επομένως, αυτό είναι ένα σημείο αναφοράς ώστε να ελέγξουμε την ορθότητα και την εγκυρότητα των αποτελεσμάτων των αλγορίθμων, αλλά και μέτρο σύγκρισης της απόδοσης τους υπολογίζοντας την απόσταση από αυτό.

2 Υλοποίηση

Για την επίλυση του προβλήματος υλοποιήσαμε τις μεθόδους steepest descent, Newton και BFGS όπου η επιλογή βήματος σε κάθε επανάληψη βασίζεται στην ευθύγραμμη αναζήτηση με ισχυρές συνθήκες Wolfe, καθώς και μία Dogleg εκδοχή της μεθόδου BFGS. Ως πρότυπο για τις υλοποιήσεις είχαμε το υλικό των διαφανειών, αλλά και το βιβλίο Numerical Optimization των J.Nocedal και S.J.Wright.

Οι υλοποιήσεις έγιναν στην γλώσσα προγραμματισμού Python, έκδοση 3.9.5. Για την αναπαράσταση του προβλήματος χρησιμοποιήθηκαν πίνακες από το πακέτο Numpy, αφού αυτοί παρέχουν καλύτερη αξιοποίηση του χώρου και οι συναρτήσεις που τους χρησιμοποιούν παρέχουν καλύτερη χρονική πολυπλοκότητα από τις προκαθορισμένες δομές της Python. Οι υπολογισμοί των μερικών παραγώγων που χρησιμοποιούν οι διάφορες μέθοδοι έχουν γίνει συμβολικά.

Όσον αφορά την ευθύγραμμη αναζήτηση με ισχυρές συνθήκες Wolfe, χρησιμοποιήσαμε τις παραμέτρους $c_1 = 10^{-4}$, $c_2 = 0.9$ και $a_{max} = 1$, ενώ η παράμετρος a_0 του αρχικού βήματος είναι ίση με 0 για την steepest descent και ίση με 1 για τις Newton και BFGS, όπως προτείνεται από την βιβλιογραφία.

Για την μέθοδο Dogleg χρησιμοποιήσαμε ως μέγιστη ακτίνα ασφαλούς περιοχής αναζήτησης την τιμή 1, αρχική ακτίνα την τιμή της μέγιστης ακτίνας και ως ελάχιστο ποσοστό βελτίωσης για αποδοχή ενός σημείου Dogleg την τιμή $\eta = 0.001$.

Για την παραγωγή των αρχικών σημείων χρησιμοποιήθηκε η γεννήτρια ψευδοτυχαίων αριθμών του πακέτου Numpy. Για να είναι εφικτή η αναπαραγωγή των αποτελεσμάτων χρησιμοποιήσαμε την τιμή 0 ως seed.

Τέλος, κάθε πείραμα τερματίζει αν ικανοποιείται κάποια από τις συνθήκες:

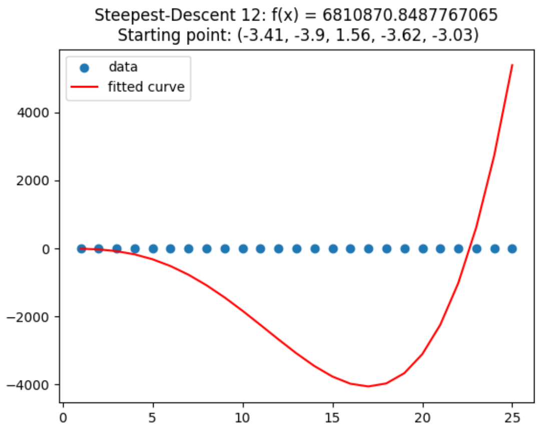
- βρέθηκε σημείο x^* με $\|\nabla f(x^*)\| \leq 10^{-8}$
- οι συναρτησιακοί υπολογισμοί της αντικειμενικής συνάρτησης υπερέβησαν τους 3000

3 Ανάλυση αλγορίθμων

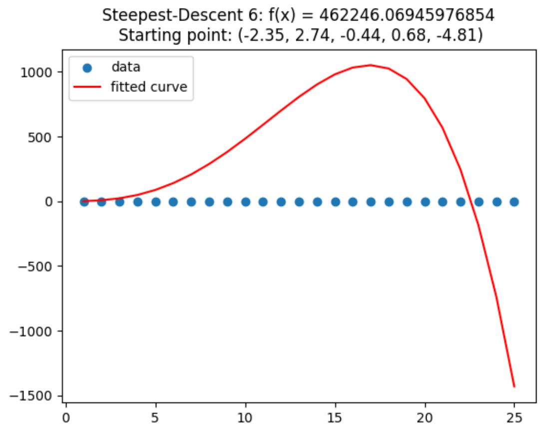
3.1 Steepest descent με Wolfe line search

Αν και από την θεωρία γνωρίζουμε ότι η μέθοδος steepest descent με line search και ισχυρές συνθήκες Wolfe τελικά συγκλίνει σε κάποιον ελαχιστοποιητή, στην πράξη κάτι τέτοιο δεν παρατηρήθηκε με επιθυμητή ακρίβεια και εντός του ορίου των συναρτησιακών υπολογισμών.

Αυτό το αποτέλεσμα δεν αποτελεί έκπληξη, καθώς είναι γνωστό ότι λόγω της φύσης του αλγορίθμου (ορθογώνια επαναληπτικά βήματα), η ταχύτητα σύγκλισης της μεθόδου είναι σχετικά μικρή. Χαρακτηριστικά, σε ένα πείραμα με μέγιστο αριθμό συναρτησιακών κλήσεων ίσο με $3 \cdot 10^5$, ο αλγόριθμος επέστρεφε ελάχιστη τιμή περίπου ίση με 0.0579, η οποία είναι πολύ μεγαλύτερη από το ολικό ελάχιστο.

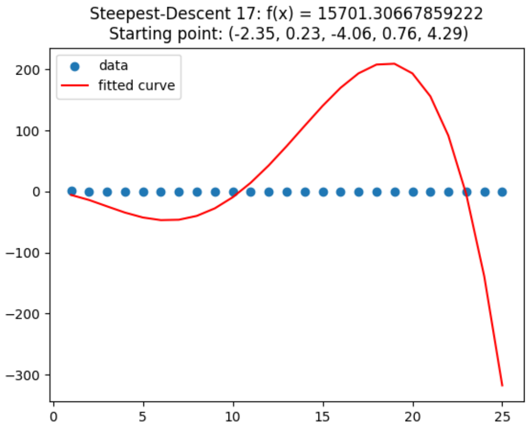


(α′) Γραφική παράσταση 1

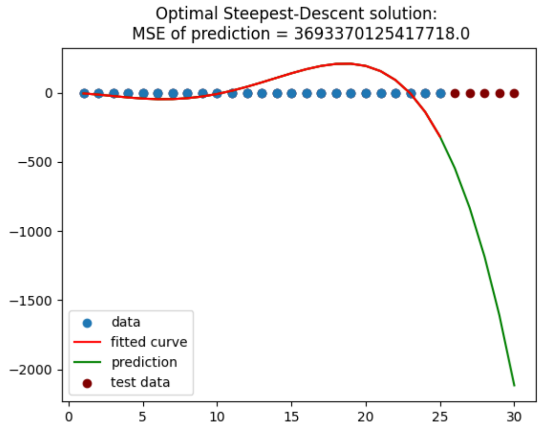


(β′) Γραφική παράσταση 2

Σχήμα 1: Οι πιο συνηθισμένες περιπτώσεις γραφικών παραστάσεων πολυωνύμων που εντοπίζει η steepest descent



(α′) Βέλτιστη λύση (θυμίζει αμυδρά την γραφική παράσταση της ολικής βέλτιστης λύσης)



(β′) Μελλοντική πρόβλεψη από βέλτιστη λύση (διαφορετική κλίμακα λόγω κακής πρόβλεψης)

Σχήμα 2: Βέλτιστη λύση και πρόβλεψη της steepest descent

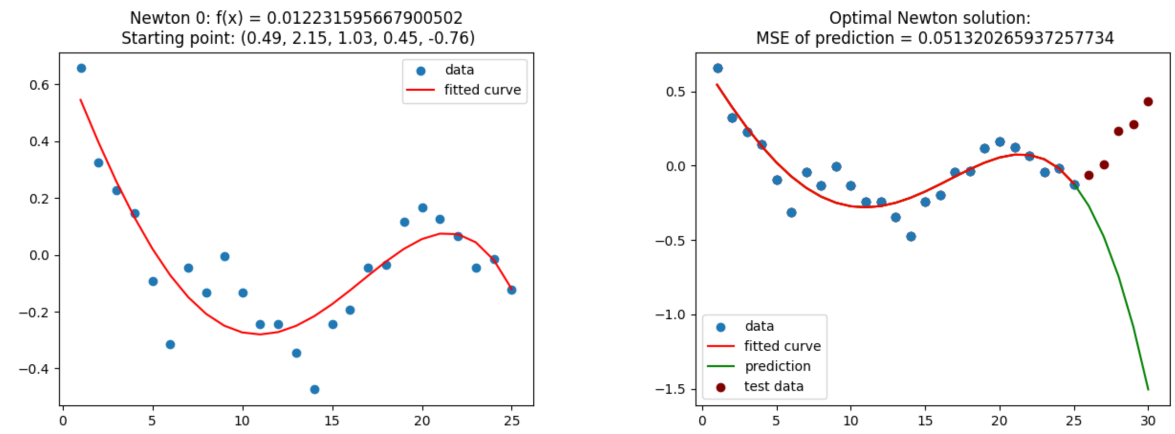
Steepest descent		
	Κλήσεις συναρτήσεων	Συναρτησιακή τιμή
mean	3022.0	5014872.136982081
median	3022.0	3434697.3918218175
min	3020	15701.30667859222
max	3025	15704385.153876271
std	1.24	4438920.463981492

Πίνακας 1: Στατιστικά μεγέθη πειραμάτων steepest descent

3.2 Newton με Wolfe line search

Αφού ο εσσιανός της αντικειμενικής συνάρτησης είναι σταθερός ως προς x, αρκεί να ελέγξουμε μόνο μία φορά στην αρχή αν είναι θετικά ορισμένος. Χρησιμοποιώντας τον αλγόριθμο του Cholesky παρατηρούμε ότι είναι θετικά ορισμένος και επομένως ο αλγόριθμος του Newton δίνει πάντα διεύθυνση καθόδου. Αυτό σε συνδυασμό με τις ισχυρές συνθήκες Wolfe εξασφαλίζει πάντα σύγκλιση, η οποία μάλιστα είναι τετραγωνική.

Πράγματι, σε όλα τα πειράματα υπήρξε σύγκλιση της μεθόδου πριν από το όριο των επαναλήψεων. Μάλιστα, όλα τα πειράματα έδωσαν σημεία πολύ κοντά στον ολικό ελαχιστοποιητή και των οποίων οι μεταξύ τους αποκλίσεις ήταν της τάξης το πολύ 10^{-16}.



(α′) Βέλτιστη λύση

(β′) Μελλοντική πρόβλεψη από βέλτιστη λύση

Σχήμα 3: Όλα τα αρχικά σημεία δίνουν σχεδόν ίδιες γραφικές παραστάσεις

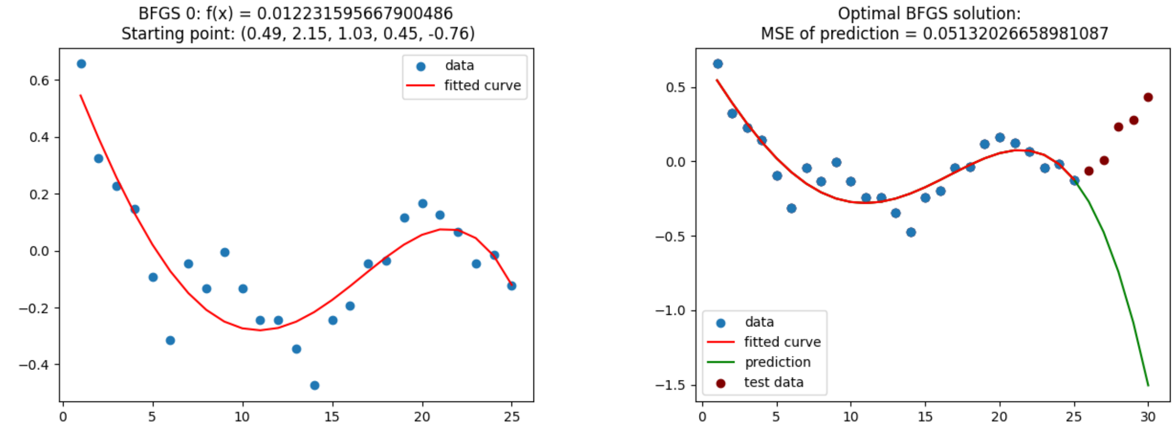
Newton		
	Κλήσεις συναρτήσεων	Συναρτησιακή τιμή
mean	58.87	0.0122315956679004958
median	47.0	0.0122315956679004988
min	7	0.0122315956679004768
max	168	0.0122315956679005078
std	53.03	$8.099505462106327e^{-18}$

Πίνακας 2: Στατιστικά μεγέθη πειραμάτων Newton

3.3 BFGS με Wolfe line search

Παρόλο που γενικά ο BFGS προτιμάται σε σχέση με τον αλγόριθμο του Newton, στην προκειμένη περίπτωση, όπου η αντικειμενική συνάρτηση είναι τετραγωνική και ο εσσιανός είναι σταθερός και θετικά ορισμένος, είναι υποδεέστερος. Έχει μεγαλύτερο υπολογιστικό κόστος ανά επανάληψη και παρουσιάζει μικρότερη ακρίβεια, αφού χρησιμοποιεί μία προσέγγιση του εσσιανού.

Σε όλα τα πειράματα υπήρξε σύγκλιση της μεθόδου πριν το όριο των επαναλήψεων και τα σημεία που παρήχθησαν ήταν πολύ κοντά στον ολικό ελαχιστοποιητή. Οι μεταξύ τους αποκλίσεις ήταν της τάξης το πολύ 10^{-12}.



(α΄) Βέλτιστη λύση

(β΄) Μελλοντική πρόβλεψη από βέλτιστη λύση

Σχήμα 4: Όλα τα αρχικά σημεία δίνουν σχεδόν ίδιες γραφικές παραστάσεις

BFGS		
	Κλήσεις συναρτήσεων	Συναρτησιακή τιμή
mean	280.53	0.012231595667900495
median	284.0	0.012231595667900497
min	185	0.012231595667900486
max	467	0.01223159566790051
std	58.54	$6.544606901286124e^{-18}$

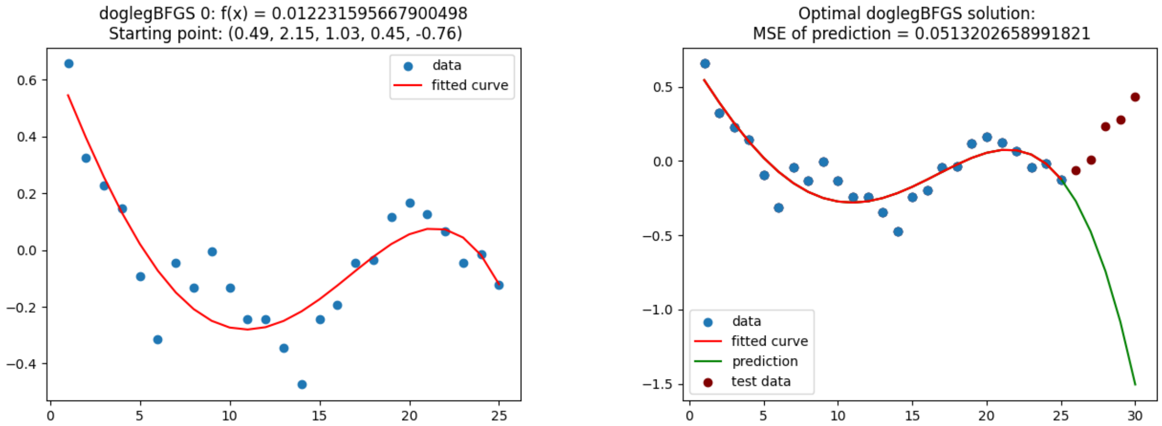
Πίνακας 3: Στατιστικά μεγέθη πειραμάτων BFGS

3.4 Dogleg BFGS

Παρατηρούμε ότι ο Dogleg BFGS βρίσκει παρόμοια αποτελέσματα σε σχέση με την line search BFGS εκδοχή του, ωστόσο είναι πολύ πιο αποδοτικός. Συγκλίνει πολύ πιο γρήγορα και με πολύ μικρότερο αριθμό κλήσεων της αντικειμενικής συνάρτησης.

Αυτό το αποτέλεσμα δεν πρέπει να μας εκπλήσσει. Εφόσον η αντικειμενική συνάρτηση είναι τετραγωνική, μία μέθοδο trust-region την προσεγγίζει πάρα πολύ καλά.

Σε όλα τα πειράματα υπήρξε σύγκλιση της μεθόδου πριν το όριο των επαναλήψεων και τα σημεία που παρήχθησαν ήταν πολύ κοντά στον ολικό ελαχιστοποιητή. Οι μεταξύ τους αποκλίσεις ήταν της τάξης το πολύ 10^{-16}.



(α΄) Βέλτιστη λύση   (β΄) Μελλοντική πρόβλεψη από βέλτιστη λύση

Σχήμα 5: Όλα τα αρχικά σημεία δίνουν σχεδόν ίδιες γραφικές παραστάσεις

Dogleg BFGS		
	Κλήσεις συναρτήσεων	Συναρτησιακή τιμή
mean	7.0	0.012231595667900495
median	6.0	0.012231595667900497
min	6	0.012231595667900486
max	36	0.01223159566790051
std	5.39	$6.381645055601919e^{-18}$

Πίνακας 4: Στατιστικά μεγέθη πειραμάτων Dogleg BFGS

4 Συμπεράσματα

Παρατηρούμε ότι η μέθοδος steepest descent παρουσιάζει την χειρότερη απόδοση τόσο σε συναρτησιακές κλήσεις, όσο και σε αποτελέσματα και προβλέψεις, αφού κάθε φορά εξαντλεί το όριο των επαναλήψεων χωρίς όμως να συγκλίνει ικανοποιητικά προς τον ελαχιστοποιητή.

Όσον αφορά τους άλλους τρεις αλγορίθμους, αυτοί παρουσιάζουν παρόμοια στατιστικά στοιχεία ως προς τα ελάχιστα που βρίσκουν. Καλύτερα αποτελέσματα επιστρέφει η Newton με απόκλιση από το ελάχιστο της τάξης 10^{-18}, ενώ οι άλλες δύο μέθοδοι έχουν σχεδόν ίδια αποτελέσματα.

Αναφορικά με την ταχύτητα σύγκλισης, η Dogleg BFGS παρουσιάζει πολύ καλύτερα αποτελέσματα από τις άλλες δύο μεθόδους, δεύτερη καλύτερη είναι η Newton και τρίτη καλύτερη είναι η BFGS. Ως προς αυτό το κριτήριο φαίνεται να υπάρχει σαφής διαχωρισμός των τεσσάρων μεθόδων.

Σύγκριση προβλέψεων	
Αλγόριθμος	Σφάλμα πρόβλεψης
Steepest descent	3693370125417718.0
Newton	0.051320265899181594
BFGS	0.05132026589918191
Dogleg BFGS	0.0513202658991821

Πίνακας 5: Σύγκριση μέσων τετραγωνικών σφαλμάτων

Όσον αφορά τις μελλοντικές προβλέψεις, όλες οι μέθοδοι, εκτός της steepest descent, παρουσιάζουν παρόμοια σφάλματα πρόβλεψης. Όπως φαίνεται και από τις γραφικές παραστάσεις, οι προβλέψεις δεν είναι ιδιαίτερα καλές. Αυτό βέβαια είναι αναμενόμενο, αφού η συμπεριφορά ενός πολυωνύμου τελικά κυριαρχείται από τον μεγιστοβάθμιό του όρο και αυτός είναι που τελικά το ωθεί προς το άπειρο. Αντίθετα, δεδομένα όπως οι τιμές ομολόγων αναμένουμε ότι θα είναι πεπερασμένες, τουλάχιστον για ένα 'κοντινό μέλλον'.

Έτσι, το μοντέλο δυσκολεύεται να εκτιμήσει trends που μπορεί να υπάρχουν στα δεδομένα. Ωστόσο, ίσως ένα μοντέλο με περισσότερα δεδομένα (για παράδειγμα μερικών ετών) θα μπορούσε να εκτιμήσει κάποιες τιμές σε ένα μικρό διάστημα του 'κοντινού μέλλοντος'. Για παράδειγμα, η τιμή της 26ης μέρας προσεγγίζεται αρκετά καλύτερα από ότι οι τιμές οι τιμές των ημερών 29 και 30.

Συνεπώς, ένα πολυωνυμικό μοντέλο δεν προσφέρει ιδιαίτερα καλές προβλέψεις. Παρόλα αυτά, ένα τέτοιο μοντέλο μπορεί να δώσει σχετικά καλές εκτιμήσεις για τις τιμές μεταξύ των παρατηρήσεων.

5 Οδηγίες για την εκτέλεση του κώδικα

Για την εγκατάσταση των πακέτων που χρησιμοποιούνται:

pip install requirements.txt

Για την εκτέλεση του κώδικα:

python main.py