



ΕΘΝΙΚΟ ΚΑΙ ΚΑΠΟΔΙΣΤΡΙΑΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ
ΤΜΗΜΑ ΜΑΘΗΜΑΤΙΚΩΝ
ΤΟΜΕΑΣ ΜΑΘΗΜΑΤΙΚΗΣ ΑΝΑΛΥΣΗΣ
ΠΑΝΕΠΙΣΤΗΜΙΟΥΠΟΛΗ, ΑΘΗΝΑ 15784
ΤΗΛ 210 - 7276397, FAX 210 - 7276398

Νούτσης Βασίλειος

Μελέτη του κανονικοποιημένου προβλήματος ελαχίστων
τετραγώνων με υπό συνθήκη λύση και εφαρμογή του στη
Βαρυτική μέθοδο

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ
ΓΙΑ ΤΟ
ΠΡΟΓΡΑΜΜΑ ΜΕΤΑΠΤΥΧΙΑΚΩΝ ΣΠΟΥΔΩΝ
ΤΟΥ ΤΜΗΜΑΤΟΣ ΜΑΘΗΜΑΤΙΚΩΝ
ΤΟΥ
ΕΘΝΙΚΟΥ ΚΑΠΟΔΙΣΤΡΙΑΚΟΥ ΠΑΝΕΠΙΣΤΗΜΙΟΥ ΑΘΗΝΩΝ
ΣΤΗΝ ΚΑΤΕΥΘΥΝΣΗ
ΕΦΑΡΜΟΣΜΕΝΑ ΜΑΘΗΜΑΤΙΚΑ

ΑΘΗΝΑ 2011

Πρόλογος

Στην παρούσα διπλωματική εργασία προσπάθησα να συνδιάσω την εμπειρία 25 χρόνων ενασχόλησης με τη Γεωφυσική και ένα κλάδο των εφαρμοσμένων μαθηματικών, την αριθμητική ανάλυση.

Επιστήμονες και μηχανικοί συχνά θέλουν να συσχετίσουν φυσικές παραμέτρους οι οποίες χαρακτηρίζουν ένα μοντέλο m και να συλλέξουν μετρήσεις (παρατηρήσεις) οι οποίες θα αποτελούν μια ομάδα (set) δεδομένων d .

Επειδή βασικές έννοιες της Φυσικής είναι γνωστές, δεχόμαστε ότι μια συνάρτηση G σχετίζει τα m και d , δηλαδή $G(m)=d$. Μια σημαντική υπόθεση είναι ότι οι μετρήσεις περιέχουν θόρυβο, ο οποίος οφείλεται στην ακρίβεια των οργάνων, δεν λαμβάνουμε υπ' όψη όλους τους παραμέτρους που επηρεάζουν το μοντέλο μας κτλ., δηλαδή θα έχουμε $G(m)=d+\eta$ όπου με η συμβολίζουμε το θόρυβο.

Το ευθύ (**forward**) πρόβλημα είναι να υπολογίσουμε τα d δοθέντος το μοντέλο m . Το αντίστροφο (**inverse**) πρόβλημα (αυτό που θα μας απασχολήσει σε αυτή την εργασία) είναι να βρούμε το μοντέλο m δοθέντος των μετρήσεων d . Στην περίπτωση ενός διακριτού γραμμικού προβλήματος έχουμε να λύσουμε το γραμμικό σύστημα $Gm=d$.

Στην παρούσα εργασία θα δούμε μερικές μεθόδους για την υπό συνθήκη λύση ενός γραμμικού συστήματος ($Ax=g$, $l_i \leq x \leq u_i$, $i=1, \dots, n$) και την εφαρμογή τους στη Γεωφυσική, στη Βαρυτική μέθοδο.

Στην Ενότητα 1 υπάρχει μια εισαγωγή γύρω από την Γεωφυσική, Βαρυτική Μέθοδος και περιγραφή του τρόπου με τον οποίο φτάνουμε από τον νόμο της βαρύτητας στον πίνακα A του γραμμικού μας συστήματος.

Στην Ενότητα 2 θα δούμε την εφαρμογή του νόμου της βαρύτητας του Νεύτωνα στην γεωφυσική και τις διορθώσεις οι οποίες χρειάζονται για να καταλήξουμε από τις μετρημένες τιμές βαρύτητας στις διορθωμένες τιμές, οι οποίες ανταποκρίνονται στις πραγματικές ανωμαλίες. Στη συνέχεια θα δούμε τον αλγόριθμο του Talwani [20], έναν τρόπο εκτίμηση του βάθους της επαφής μεταξύ δύο στρωμάτων με διαφορετική πυκνότητα.

Στην Ενότητα 3 θα δούμε τον τύπο εκτίμησης του σφάλματος μεταξύ των διορθωμένων τιμών βαρύτητας (g) και της ανταπόκρισης κάποιου μοντέλου (κάποια προσέγγιση g^*). Σύγκριση μεταξύ του τύπου σφάλματος που μας δίνει η νόρμα των υπολοίπων $\|g-g^*\|$ (όπου με $\|\cdot\|$ θα συμβολίζουμε την Ευκλείδεια νόρμα) με τον τύπο των C. Brezinski, G. Rodriguez, S. Seatsu [2, 3, 4].

Στην Ενότητα 4 θα δούμε 4 μεθόδους για τη λύση του προβλήματος.

- Τον επαναληπτικό αλγόριθμο του Kaczmarz [8] για λύση γραμμικού συστήματος και προσαρμογή του στη λύση γραμμικού συστήματος με υπό συνθήκη λύση.
- Τον επαναληπτικό αλγόριθμο των Levenberg-Marquardt [5, 6] και τη μετατροπή του σε Modified LM για τη λύση των γραμμικών συστημάτων με υπό συνθήκη λύση.
- Την εφαρμογή της κανονικοποίησης Tichonov και επιλογή της κατάλληλης παραμέτρου λ . Για τη λύση του συστήματος χρησιμοποιούμε τις μεθόδους SVD, GSVD.

Αυτή η μέθοδος είναι η μόνη που λαμβάνει υπόψη ότι οι μετρήσεις μας g περιέχουν θόρυβο. Ιδιαίτερο ενδιαφέρον έχει η μέθοδος παρεμβολής που την προτείνει ο κύριος Brezinski στην εργασία [19]

- Λύση της MLM με SVD. Συνδυασμός των δύο προηγούμενων μεθόδων.

Τα αποτελέσματα των δοκιμών θα τα δούμε στην Ενότητα 5.

Στην Ενότητα 6 παίρνουμε το αποτέλεσμα, την Γεωλογική Ερμηνεία.

Ευχαριστώ την επιβλέπουσα καθηγήτρια κυρία Μ. Μητρούλη και τα μέλη της επιτροπής καθηγητές κυρίους Β. Δουγκαλή και Σ. Νοτάρη για τις πολύτιμες συμβουλές και παρατηρήσεις τους που συνέλαβαν στην ολοκλήρωση της παρούσας εργασίας.

Περιεχόμενα

1	Εισαγωγή	5
2	Η βαρυτική μέθοδος	6
2.1	Η επίδραση του σχήματος της Γης	7
2.2	Διορθώσεις βαρύτητας	7
2.2.1	Γεωγραφικό πλάτος	8
2.2.2	Υψόμετρο	8
2.2.3	Υλικό που παρεμβάλετε μεταξύ των σημείων μέτρησης και αναφοράς	9
2.2.4	Τοπογραφία	9
2.2.5	Σεληνιακή διόρθωση	9
2.3	Βαρυτική έλξη. Εξίσωση πρίσματος	10
2.4	Εφαρμογή	10
2.5	Μέθοδος καταγραφής μετρήσεων	11
2.6	Μέθοδος Talwani.	14
2.7	Ανάγκη χρήσης μεθόδων Αριθμητικής Ανάλυσης	17
2.7.1	Επιλογή Μοντέλου	17
3	Εκτίμηση σφάλματος	18
3.1	Εισαγωγή	18
3.2	Εκτίμηση του σφάλματος κατά Brezinski	18
4	Επαναληπτικές μέθοδοι εκτίμησης παραμέτρων με υπό συνθήκη λύση	23
4.1	Αλγόριθμος Kaczmarz. Προσαρμογή του για λύση του προβλήματος	23
4.2	Αλγόριθμος Levenberg-Marquardt	26
4.3	Προσαρμογή του αλγορίθμου LM σε ModifiedLM	27
4.4	Μέθοδοι Κανονικοποίησης και Κανονικοποίηση Tichonov.	30
4.5	Τεχνικές παρεμβολής για γραμμικά συστήματα σε κακή κατάσταση	35
4.6	Λύση του MLM με SVD	36
5	Αριθμητικά Αποτελέσματα	39
5.1	Εκτίμηση σφάλματος	39
5.2	Συνάρτηση LSQLIN της Matlab	45
5.3	Αλγόριθμος InvKaczmarz	46
5.4	Αλγόριθμος ModifiedLM	48
5.5	Κανονικοποίηση Tichonov	49
5.6	Τεχνικές παρεμβολής για γραμμικά συστήματα σε κακή κατάσταση	53
5.7	Λύση MLM με SVD	53
6	Γεωλογική Ερμηνεία	55

1 Εισαγωγή

Στην Γεωφυσική εφαρμόζουμε τους νόμους της Φυσικής στο να εκτιμήσουμε κάποιες από τις ιδιότητες των πετρωμάτων σε μια περιοχή έρευνας. Αυτές οι ιδιότητες μπορούν να είναι πυκνότητα, μαγνητική επιδεκτικότητα, ειδική ηλεκτρική αντίσταση, κτλ. Εκτιμάμε την κατανομή αυτών των ιδιοτήτων στο χώρο και στη συνέχεια μελετάμε την αντιστοίχσή τους με γεωλογικούς σχηματισμούς.

Η Βαρυτική μέθοδος χρησιμοποιείται ευρέως στη γεωφυσική για την αναγνώριση επαφών μεταξύ δύο σχηματισμών με διαφορετική πυκνότητα.

Η διαδικασία είναι η ακόλουθη: οι μετρήσεις στην επιφάνεια του εδάφους γίνονται σε κানাβο ή σε ευθύγραμμη τομή. Οι μετρήσεις σε κানাβο μας επιτρέπουν μια τρισδιάστατη ερμηνεία ενώ οι μετρήσεις σε τομές μια δυσδιάστατη.

Στην παρούσα εργασία χρησιμοποιούμε καταρχάς μια συγκεκριμένη μέθοδο δυσδιάστατης ερμηνείας που προτάθηκε από τον Talwani [20] και το εφαρμόζουμε σε τομές. Το αποτέλεσμα είναι τα σημεία επαφής ανάμεσα σε δύο στρώματα με διαφορετικές πυκνότητες: το επιφανειακό στρώμα με μικρότερη πυκνότητα (συνήθως) και το βάθος στο οποίο εμφανίζεται το κάτω στρώμα με διαφορετική πυκνότητα. Το μειονέκτημα της μεθόδου είναι ότι χρησιμοποιεί μόνο δύο τιμές πυκνότητας στην ερμηνεία: όλα τα πετρώματα που βρίσκονται στο άνω στρώμα τα προσεγγίζουμε με μια τιμή πυκνότητας και όλα τα πετρώματα στο κάτω στρώμα με μια άλλη. Στη φύση δεν υπάρχει τόσο απότομη μεταβολή στις πυκνότητες στο σημείο επαφής των δύο στρωμάτων, η μεταβολή γίνεται σταδιακά. Επίσης, δεν μπορούμε να περιγράψουμε ένα έγκοιλο (κενό, σπηλιά) μέσα στην Γη, άρα πρέπει να χρησιμοποιήσουμε διαφορετική περιγραφή για το μοντέλο.

Προσεγγίζοντας την κατανομή πυκνότητας του υπεδάφους με σύνολο n οριζόντιων πρισμάτων άπειρου μήκους και διαφορετικών πυκνοτήτων καταλήγουμε σε ένα πρόβλημα αντιστροφής. Το ζητούμενο είναι να βρούμε αυτές τις πυκνότητες κατά τρόπο ώστε η βαρυτική επίδραση του μοντέλου να προσεγγίζει κατά τον βέλτιστο τρόπο το σύνολο των μετρήσεων. Χρησιμοποιώντας τον τύπο του Banerjee[1] υπολογίζουμε ένα γεωμετρικό παράγοντα (geometric factor) A_{ij} ανάμεσα σε κάποιο σημείο i της τομής και κάποιου πρίσματος j του μοντέλου. Λαμβάνοντας υπόψη όλα τα σημεία της τομής και όλα τα πρίσματα του μοντέλου δημιουργούμε τον πίνακα $A \in R^{n \times n}$ και μαζί με μια αρχική εκτίμηση των πυκνοτήτων των πρισμάτων x_i^0 , $i=1, \dots, n$ παίρνουμε την ανταπόκριση του μοντέλου $\tilde{g} = Ax^0$. Το σημαντικότερο είναι ότι ο τετραγωνικός πίνακας A έχει πολύ κακή κατάσταση (ill conditioned). Για ένα μοντέλο με 13 σειρές πρισμάτων, σύνολο 300 πρίσματα, ο πίνακας A έχει δείκτη κατάστασης (condition number) $\kappa(A) := \|A\| \cdot \|A^{-1}\| = 8.4 \times 10^{12}$. Αυτό είναι ένα δύσκολο αριθμητικό πρόβλημα, το οποίο χρειάζεται ιδιαίτερη μεταχείριση, διαφορετικά εισάγει μεγάλα σφάλματα στους υπολογισμούς. Επιπλέον, οι πυκνότητες των πρισμάτων πρέπει να κυμαίνονται μεταξύ $2.0-2.7 \times 10^3 \text{ kg} \times \text{m}^{-3}$ άρα σκοπός μας είναι να βρούμε το $\min_{\mathbf{l}_i \leq \mathbf{x}_i \leq \mathbf{u}_i} \|\mathbf{Ax} - \mathbf{g}\|$, $\mathbf{i} = \mathbf{1}, \dots, \mathbf{n}$ όπου τα \mathbf{l}_i , \mathbf{u}_i είναι τα κάτω, άνω όρια των πυκνοτήτων.

2 Η βαρυτική μέθοδος

Όλες οι βαρυτικές μελέτες στηρίζονται στο νόμο βαρύτητας του Νεύτωνα, ο οποίος εκφράζεται ως η δύναμη F μεταξύ δύο σημειακών μαζών m_1 και m_2 σε απόσταση r και δίνεται από τη σχέση

$$F = \frac{Gm_1m_2}{r^2} \quad \text{σε Newtons} \quad (2.1)$$

όπου $G = 6.67 \times 10^{-11} \text{N} \cdot \text{m}^2 \cdot \text{kg}^{-2}$ είναι η παγκόσμια σταθερά της βαρύτητας.

Στην περίπτωση της έλξης μιας μάζας m από τη Γη (μάζα M), η μάζα m πέφτει κατακόρυφα με μια επιτάχυνση $g \approx 9.8 \text{m} \cdot \text{s}^{-2}$. Η επιτάχυνση η οποία οφείλεται στη βαρύτητα δεν είναι ακριβώς η ίδια σε όλα τα σημεία της επιφάνειας της Γης. Υπάρχουν μικρές αποκλίσεις που προκύπτουν από μεταβολές του σχήματος της Γης και από ανομοιογένειες της πυκνότητας αυτής. Γεωφυσικό ενδιαφέρον προκύπτει για τη μέτρηση και ερμηνεία των μικρών μεταβολών του g , σχετικά με το σχήμα και τη δομή της Γης.

Η δύναμη που ασκείται σε μια πύπτουσα μάζα m προς τη Γη υπό την επίδραση μόνο της βαρύτητας γράφεται :

$$F = \text{μάζα} \times \text{επιτάχυνση} = mg \quad (2.2)$$

Συνδυάζοντας την 2.1 με την 2.2 προκύπτει

$$F = mg = \frac{GmM}{r^2} \quad (2.3)$$

όπου r είναι η απόσταση μέχρι το κέντρο της Γης. Από 2.2 και 2.3 έχουμε:

$$g = \frac{GM}{r^2} \quad (2.4)$$

Η γενική σχέση μεταξύ της βαρύτητας και της μέσης πυκνότητας της Γης ρ προσδιορίζεται εύκολα: αν θεωρούμε τη Γη σφαιρική προκύπτει

$$\rho = \frac{\text{μάζα}}{\text{όγκος}} = \frac{M}{\frac{4}{3}\pi r^3} = \frac{3M}{4\pi r^3} \quad (2.5)$$

Από την 2.4 προκύπτει $M = \frac{gr^2}{G}$ άρα

$$\rho = \frac{3g}{4\pi rG} \quad (2.6)$$

Χρησιμοποιώντας πρόσφατες μετρήσεις και την εξίσωση 2.6, η μέση πυκνότητα της Γης υπολογίστηκε σε $5.52 \times 10^3 \text{kg} \cdot \text{m}^{-3}$. Επειδή αυτή η τιμή είναι πολύ μεγαλύτερη από τις πυκνότητες των επιφανειακών πετρωμάτων ($2.0\text{--}3.0 \times 10^3 \text{kg} \cdot \text{m}^{-3}$) συμπεραίνουμε ότι η πυκνότητα της Γης πρέπει να αυξάνεται προς το κέντρο της.

Ο γεωφυσικός ενδιαφέρεται για τη χρησιμοποίηση του πεδίου βαρύτητας της Γης για να μελετήσει το σχήμα ή/και τη δομή της. Χρησιμοποιεί την σχέση 2.6 με κάποια αναδιάθρωση των όρων της:

$$g = \frac{4\pi\rho rG}{3} \quad (2.7)$$

Εάν η Γη ήταν τελείως σφαιρική και τελείως ομογενής (ρ, r - σταθερές) το g θα είχε ίδια τιμή παντού στην επιφάνεια της. Εν τούτοις, εάν σε κάποιο σημείο της Γης υπάρχει απόκλιση από τη σφαιρικότητα ή υπάρχει μια τοπική ανωμαλία στην πυκνότητα, το g θα αποκλίνει.

2.1 Η επίδραση του σχήματος της Γης

Επειδή υπάρχει η πλάτυνση της Γης, δηλαδή η ακτίνα της δεν είναι παντού η ίδια, το g μεταβάλλεται με το γεωγραφικό πλάτος. Σύμφωνα με τη **Διεθνή Εξίσωση Βαρύτητας (Geodetic Reference System 1967)** που έχει υιοθετηθεί από τη **Διεθνή Ένωση Γεωδαισίας και Γεωφυσικής**, η τιμή του g στο επίπεδο της θάλασσας μεταβάλλεται με το πλάτος (ϕ) σύμφωνα με τη σχέση :

$$g_t = 978031.85(1 + 0.00527889 \sin^2(\phi) + 0.000023462 \sin^4(\phi)) \quad \text{mGal} \quad (2.8)$$

Σε αυτή την εξίσωση το ϕ μετριέται σε μοίρες. Η τιμή 978031.85 mGal είναι η τιμή του g_t στον ισημερινό ($\phi = 0$). Καθώς το ϕ αυξάνει προς τους πόλους το g θα αυξάνει επειδή η απόσταση μεταξύ ενός σημείου που βρίσκεται επί της στάθμης της θάλασσας και του κέντρου της Γης μειώνεται. Στους πόλους προκύπτει $g_t = 983217.72 \text{mGal}$. Η βαρύτητα μεταβάλλεται περίπου 5000 mGal από τον ισημερινό μέχρι τους πόλους. Επίσης, η φυγοκεντρική επίδραση, η οποία δρα σε κάθε αντικείμενο που επικάθεται στη Γη, πρέπει να ληφθεί υπόψη. Η ένταση της φυγοκεντρικής δύναμης είναι:

$$\alpha_c = \omega^2 d \quad (2.9)$$

όπου ω είναι η γωνιακή ταχύτητα περιστροφής και d είναι η απόσταση του αντικειμένου από το κέντρο της κυκλικής τροχιάς. Βλέπουμε ότι η φυγοκεντρική επίδραση είναι μεγαλύτερη στον ισημερινό επειδή εκεί η τιμή της απόστασης d είναι η μεγαλύτερη.

Από τα παραπάνω είναι φανερό γιατί η βαρύτητα δεν είναι παντού η ίδια σε ένα ελλειψοειδές εκ περιστροφής. Η γενική εξίσωση για τον υπολογισμό της βαρύτητας $g(\phi)$ για ένα πλάτος ϕ σε ένα ελλειψοειδές εκ περιστροφής είναι

$$g(\phi) = g_e(1 + C_1 \sin^2(\phi) + C_2 \sin^4(\phi)) \quad (2.10)$$

Οι σταθερές C_1, C_2 εξαρτώνται από την πλάτυνση και την ταχύτητα περιστροφής, υπολογίζονται από αστρονομικές μετρήσεις και παρατηρήσεις των τροχιών τεχνητών δορυφόρων.

2.2 Διορθώσεις βαρύτητας

Γενικά, μετριέται το g σε ένα σημείο στη περιοχή έρευνας και σε ένα σημείο αναφοράς μακριά από την περιοχή, αλλά ερμηνεύεται η τιμή Δg , δηλαδή η διαφορά μεταξύ της βαρύτητας στο σημείο μέτρησης και της βαρύτητας στο σημείο αναφοράς. Επειδή τέτοιες ανωμαλίες βαρύτητας είναι πολύ μικρές, η μονάδα $\text{m} \cdot \text{s}^{-2}$ είναι υπερβολικά μεγάλη (κυμαίνεται μεταξύ 9.78 και 9.83 $\text{m} \cdot \text{s}^{-2}$).

Οι ανωμαλίες βαρύτητας μετρούνται σε μονάδες βαρύτητας (g.u.) όπου :

$$1\text{Gal} = 1\text{cm} \cdot \text{s}^{-2} = 0.01\text{m} \cdot \text{s}^{-2} \quad \text{ή}$$

$$1\text{mGal} = 10^{-3}\text{Gal} = 10^{-3}\text{cm} \cdot \text{s}^{-2} = 10^{-5}\text{m} \cdot \text{s}^{-2} = 10\text{g.u.}$$

Από την στιγμή που έχουμε τις μετρήσεις, στη συνέχεια χρειάζεται να εφαρμοσθούν “διορθώσεις” για να απαλειφθούν οι άλλες επιδράσεις, πριν κανείς προχωρήσει και συμπεράνει για τη φυσική κατάσταση του φλοιού στην περιοχή της έρευνας. Παρακάτω αναφέρονται οι κυριότερες διορθώσεις για να πάμε από τις μετρημένες τιμές στις διορθωμένες τιμές βαρύτητας.

2.2.1 Γεωγραφικό πλάτος

Έχω αναφέρει ότι η βαρύτητα μεταβάλλεται με το γεωγραφικό πλάτος. Η επίδραση του πλάτους πρέπει να εξαλειφθεί από τη μετρηθείσα τιμή, πριν δοθεί οποιαδήποτε ερμηνεία σχετικά με τις μεταβολές της πυκνότητας.

$$\Delta_1(g) = 978031.85(1 + 0.00527889 \sin^2(\phi) + 0.000023462 \sin^4(\phi)) \text{mGal} \quad (2.11)$$

2.2.2 Υψόμετρο

Η εξίσωση 2.4 δείχνει ότι το g μεταβάλλεται με την απόσταση r από το κέντρο της Γης σύμφωνα με το λόγο $1/r^2$. Αυτό σημαίνει ότι το g ελαττώνεται με την αύξηση του υψόμετρου από την επιφάνεια της Γης και επομένως, εάν τα σημεία μέτρησης και αναφοράς δεν είναι στο ίδιο υψόμετρο, τότε προκύπτει μια επίδραση στο Δg .

Ας υποθέσουμε ότι το σημείο μέτρησης S βρίσκεται σε ύψος h πάνω από το σημείο αναφοράς. Εάν το σημείο αναφοράς βρίσκεται σε απόσταση r από το κέντρο της Γης, το σημείο μέτρησης θα είναι σε μια απόσταση $(r+h)$. Εάν g_p είναι η τιμή της βαρύτητας στο σημείο αναφοράς τότε από τη σχέση 2.4 έχουμε :

$$g_p = \frac{GM}{r^2} \quad (2.12)$$

Εάν g_s είναι η τιμή της βαρύτητας στο σημείο S , τότε έχουμε

$$g_s = \frac{GM}{(r+h)^2} \quad (2.13)$$

διαιρώντας την εξίσωση 2.13 με την 2.12 έχουμε:

$$\frac{g_s}{g_p} = \frac{r^2}{(r+h)^2} \quad (2.14)$$

Εάν το υψόμετρο h είναι μικρό, αναπτύσσοντας σε διωνυμική σειρά και λαμβάνοντας μόνο τους δύο πρώτους όρους έχουμε :

$$g_s = g_p \left(1 - \frac{2h}{r}\right) = g_p - \frac{2hg_p}{r} \quad (2.15)$$

Η διαφορά βαρύτητας $\Delta_2 g$ μεταξύ των σημείων μέτρησης και αναφοράς που οφείλεται στην επίδραση του υψόμετρου είναι :

$$\Delta_2 g = g_p - g_s = \frac{2hg_p}{r} \quad (2.16)$$

Λαμβάνοντας ως r την μέση ακτίνα της Γης ($R= 6367 \text{ km}$) και ως μέση τιμή του $g_p = 980625 \text{mGal}$, η σχέση 2.16 γίνεται :

$$\Delta_2 g = 3.086 \cdot h \text{ g.u.} \quad (2.17)$$

όπου το υψόμετρο h μετριέται σε μέτρα.

Αυτή η διόρθωση είναι γνωστή ως **διόρθωση ελεύθερου αέρα**.

2.2.3 Υλικό που παρεμβάλετε μεταξύ των σημείων μέτρησης και αναφοράς

Πέραν της υψομετρικής διαφοράς μεταξύ των σημείων μέτρησης και αναφοράς που οφείλεται στην τοπογραφία, υπάρχει και η μάζα που δεν έχει ληφθεί υπόψη. Αν το σημείο μέτρησης βρίσκεται σε υψόμετρο h κάτω από το σημείο αναφοράς, το Δ_2g θα είναι αρνητικό. Την ίδια στιγμή, η βαρύτητα στο σημείο μέτρησης θα είναι μεγαλύτερη από εκείνη στο σημείο αναφοράς κατά ένα ποσό Δ_3g , επειδή ασκείται μια επιπλέον έλξη από τη μάζα πυκνότητας ρ , που βρίσκεται μεταξύ των επιπέδων του σημείου μέτρησης και του σημείου αναφοράς. Μπορεί ναδειχθεί ότι:

$$\Delta_3g = 2\pi G\rho h = 41.91 \times 10^{-5} \cdot \rho \cdot h \text{ g.u.} \quad (2.18)$$

όπου h δίνεται σε μέτρα και το ρ σε kg m^{-3} .

Η διόρθωση Δ_3g αφαιρείται από το Δg εάν το σημείο μέτρησης είναι υψηλότερα από το σημείο αναφοράς και προστίθεται στο Δg εάν το σημείο μέτρησης είναι χαμηλότερα του σημείου αναφοράς. Η ποσότητα Δ_3g είναι γνωστή ως διόρθωση Bouguer και λαμβάνει υπόψη, αντίθετα με τη διόρθωση ελεύθερου αέρα, και την πυκνότητα ρ .

2.2.4 Τοπογραφία

Ένας λόφος πλησίον στο σημείο μέτρησης ασκεί μια έλξη της οποίας το αποτέλεσμα θα έχει φορά προς τα πάνω. Επίσης, μια κοιλάδα πλησίον στο σημείο μέτρησης, αφαιρεί μια ελκτική μάζα και στο σημείο μέτρησης έχουμε χαμηλότερη Δg από την πραγματική.

Σύμφωνα με τα παραπάνω προκύπτει ότι, η μέτρηση βαρύτητας πρέπει να διορθωθεί από επιδράσεις λόφων και κοιλάδων που βρίσκονται κοντά στο σημείο μέτρησης. Ουσιαστικά, πρέπει να ληφθεί υπόψη η επίδραση της τοπογραφίας. Η επίδραση αυτή υπολογίζεται διαιρώντας την περιοχή γύρω από το σημείο μέτρησης σε "διαμερίσματα" (κάναβος με υψόμετρα).

Το μέσο υψόμετρο καθενός διαμερίσματος υπολογίζεται, αδιαφορώντας για το πρόσημο του υψόμετρου επειδή η διόρθωση είναι πάντα προσθετική (κάθε μεταβολή στη τοπογραφία ελαττώνει την βαρύτητα και κατά συνέπεια η διόρθωση πρέπει να είναι πάντα προσθετική στο Δg).

Η τοπογραφική διόρθωση Δ_4g η οποία πρέπει να προστεθεί στο Δg λαμβάνεται από το άθροισμα των διορθώσεων των επιμέρους διαμερισμάτων. Έτσι βρίσκουμε μια τιμή η οποία έχει απαλλαχθεί από όλες τις επιδράσεις, εκτός από τις μεταβολές της πυκνότητας κάτω από την επιφάνεια του εδάφους στην περιοχή της έρευνας.

$$\text{Διορθωμένη τιμή} = \Delta g \pm \Delta_1g \pm \Delta_2g \pm \Delta_3g + \Delta_4g \quad (2.19)$$

2.2.5 Σεληνιακή διόρθωση

Επίσης πρέπει να αφαιρεθεί από τις μετρήσεις η έλξη της Σελήνης. Τα καινούργια όργανα έχουν τη δυνατότητα να την αφαιρέσουν: έχουν αποθηκευμένη την ημερομηνία, τις συντεταγμένες του σημείου μέτρησης και την ώρα και μπορούν να υπολογίσουν την σεληνιακή διόρθωση.

Επειδή ο Ήλιος και οι υπόλοιποι πλανήτες είναι αρκετά μακριά, η βαρυτική τους επιρροή θεωρείται σταθερή σε όλη την επιφάνεια της Γης άρα προσθέτει μια σταθερή ποσότητα σε όλες μας τις μετρήσεις. Επειδή αυτό που μας ενδιαφέρει είναι οι διαφορές της βαρύτητας από ένα σημείο στο άλλο, δεν είναι ανάγκη να αφαιρεθεί από τις μετρήσεις.

Μετά τις διορθώσεις, οι τιμές βαρύτητας έχουν αναχθεί στο επίπεδο τις θάλασσας.

2.3 Βαρυτική έλξη. Εξίσωση πρίσματος

Η κάθετη συνιστώσα της βαρυτικής έλξης ενός πρίσματος δεσμευμένο από τα επίπεδα

$$x = x_1, x = x_2 : y = y_1, y = y_2 : z = z_1, z = z_2$$

δίνεται από τον τύπο:

$$g_z = G\rho \int_{x_1}^{x_2} \int_{y_1}^{y_2} \int_{z_1}^{z_2} \frac{z dx dy dz}{(x^2 + y^2 + z^2)^{3/2}}$$

όπου G είναι η παγκόσμια σταθερά της βαρύτητας και ρ είναι η πυκνότητα του πρίσματος. Στο άρθρο [1] ο Banerjee καταλήγει στον τύπο:

$$g_z = G\rho \left[x \ln(y + \sqrt{x^2 + y^2 + z^2}) + y \ln(x + \sqrt{x^2 + y^2 + z^2}) - z \arctan \frac{xy}{z\sqrt{x^2 + y^2 + z^2}} \right] \Bigg|_{x_1}^{x_2} \Bigg|_{y_1}^{y_2} \Bigg|_{z_1}^{z_2}$$

Στο ίδιο άρθρο βλέπουμε ότι αν θέλουμε να χρησιμοποιήσουμε τον τύπο για δυσδιάστατη απεικόνιση (τομογραφία) απειρίζουμε τη y διάσταση.

Η συνάρτηση **fCalcA**(x_1, x_2, y_1, y_2) υπολογίζει αυτόν τον γεωμετρικό παράγοντα για ένα πρίσμα με συντεταγμένες x_1, x_2, z_1, z_2 σε σχέση με ένα σημείο μέτρησης με συντεταγμένες xms, zms .

2.4 Εφαρμογή

Στην περιοχή του Θριάσιου Πεδίου, το Ινστιτούτο Γεωλογικών και Μεταλλευτικών Ερευνών (IGME) έκανε μια γεωλογική έρευνα για την αποτύπωση των επιφανειακών στρωμάτων και την εκτίμηση των υδατικών αποθεμάτων, έρευνα στην οποία συμμετείχα και εγώ.

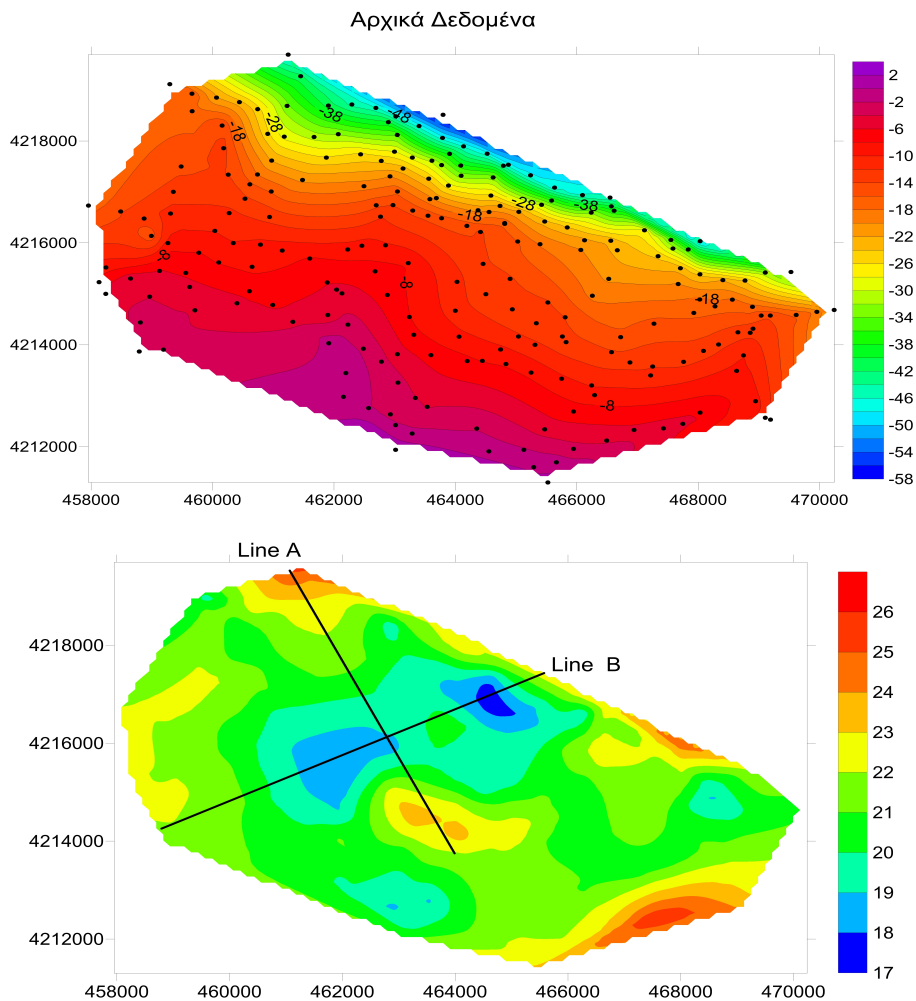
Ως σημείο αναφοράς χρησιμοποιήθηκε μια βαρυτική βάση της Γεωγραφικής Υπηρεσίας Στρατού κοντά στην περιοχή της έρευνας. Σκοπός μας είναι να καλύψουμε με μετρήσεις όλη την περιοχή έρευνας. Στις μετρήσεις της βαρύτητας χρησιμοποιούμε το mGal (ενότητα 2.2).

Μετρήσαμε της συντεταγμένες κάθε σημείου (οι μαύρες κουκίδες στον κάτω χάρτη) με διαφορικό GPS διότι θέλαμε μεγάλη ακρίβεια στο υψόμετρο: σύμφωνα με την εξίσωση 2.17 τα 10cm σφάλμα στο υψόμετρο μας δίνει μια διαφορά 0.3 mGal στη τιμή της Δg , σφάλμα αρκετά μεγάλο. Τις συντεταγμένες κάθε σημείου τις μετέτρεψα στο προβολικό σύστημα ΕΓΣΑ 87 (Ελληνικό Γεωδαιτικό Σύστημα Αναφοράς) διότι είναι το πλέον πρόσφατο προβολικό σύστημα που χρησιμοποιείται στην Ελλάδα. Αποτέλεσμα αυτού, κάθε συντεταγμένη μετατρέπεται από μοίρες σε μέτρα.

Εφαρμόζοντας τις παραπάνω διορθώσεις $\Delta_1 g, \Delta_2 g, \Delta_3 g, \Delta_4 g$ καταλήγουμε στον χάρτη :

Για την τοπογραφία χρησιμοποίησα ψηφιοποιημένους χάρτες της Γεωγραφικής Υπηρεσίας Στρατού. Καταλήγουμε σε ένα κανονικό κάρναβο με απόσταση 200 μέτρα.

Κάθε μέρα, πριν αρχίσουμε τις μετρήσεις, μετράμε στο σταθμό βάσης. Το ίδιο γίνεται και στο τέλος της ημέρας. Από της μετρήσεις αφαιρούμε την κλίση η οποία οφείλεται στην ολίσθηση του ελατηρίου του οργάνου.

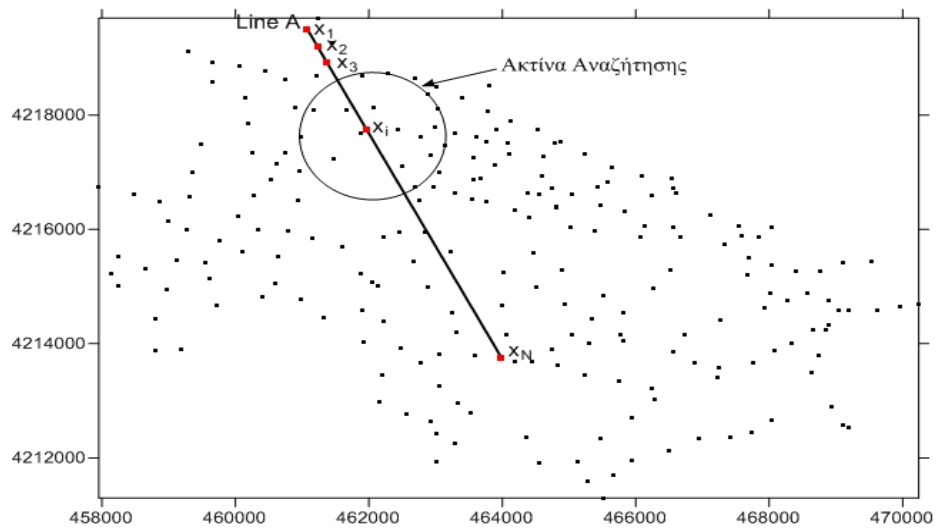


Με μια πρώτη ματιά, το γεωλογικό μοντέλο πάνω στην γραμμή A ερμηνεύεται με μια "λεκάνη", βορειοανατολικά και νοτιοδυτικά υψηλές τιμές (τα καφέ-κόκκινα χρώματα που αντιστοιχούν σε υψηλές τιμές βαρύτητας δηλώνουν ότι το υπόβαθρο, με πυκνότητα 2.50-2.70 βρίσκεται πιο κοντά στην επιφάνεια) και στις χαμηλότερες τιμές στη μέση της γραμμής (τα σιέλ-μπλε χρώματα που αντιστοιχούν στις χαμηλές τιμές βαρύτητας μας δείχνουν ότι το υπόβαθρο βρίσκεται πιο βαθιά, άρα οι μεγάλες πυκνότητες επηρεάζουν λιγότερο τις μετρήσεις). Πάνω από το υπόβαθρο έχουμε επιφανειακά ιζήματα με πυκνότητα 2.10-2.40. Να δούμε όμως τι θα δείξουν και τα αποτελέσματα της ερευνάς μας.

2.5 Μέθοδος καταγραφής μετρήσεων

Η γραμμή A έχει μήκος περίπου 6500 μέτρα. Με ένα αλγόριθμο **Inverse Distance** παίρνουμε μετρήσεις κάθε 100 μέτρα, στα σημεία x_1, \dots, x_N κατά μήκος της γραμμής. Χρησιμοποιώ ακτίνα αναζήτησης 500 μέτρα (για ακτίνα αναζήτησης π.χ. 200 θα έχουμε πιο απότομες αλλαγές, η καμπύλη θα είναι λιγότερο ομαλή). Με τον ίδιο αλγόριθμο έχουμε και τα υψόμετρα κάθε 100 μέτρα.

Για κάποιο σημείο x_i του παραπάνω σχήματος, όλα τα μετρημένα σημεία (οι μικρές μαύρες κου-



κίδες) μέσα στην ακτίνα αναζήτησης επηρεάζουν το σημείο μας με μια δύναμη ανάλογη του αντίστροφου της απόστασής τους. Αποτέλεσμα αυτού, μπορούμε να τοποθετήσουμε την τομή μας στην αρχή των αξόνων, κατά μήκος του X-άξονα και ως Y-άξονα το βάθος της ερμηνείας μας.

Ξεκινήσαμε από τις μετρημένες τιμές g_m , μετά τις διορθώσεις πήραμε τις g_B (Bouguer) καταλήγοντας μέσω του αλγορίθμου InversDistance στα $x_{p_i}, g_{p_i}, z_{p_i}, i = 1, \dots, n$ τα οποία αντιπροσωπεύουν την απόσταση από την αρχή της τομής, την βαρύτητα και το υψόμετρο κάθε σημείου.

Αλγόριθμος InversDistance

Input: $x_{bi}, y_{bi}, z_{bi}, g_{bi}, i = 1, \dots, n$ % Συντεταγμένες (x,y), υψόμετρο (z), τιμή βαρύτητας (g), για τα n μετρημένα σημεία
 $x1, x2, y1, y2$ % Συντεταγμένες της τομής, (x1, y1) για την αρχή της και (x2, y2) για το τέλος της.
 $rint, stp$ % Εσωτερική ακτίνα αναζήτησης, Βήμα
Output: $x_{pi}, z_{pi}, g_{pi}, i = 1, \dots, n$ % Απόσταση από την αρχή της τομής, υψόμετρο και τιμή βαρύτητας για κάθε σημείο

Begin

$dy = y2 - y1, dx = x2 - x1$

$dst = \sqrt{dx^2 + dy^2}$ % Υπολογισμός μήκους τομής

$ang = \arctan\left(\frac{dy}{dx}\right)$ % Υπολογισμός κλήσης τομής

$n = \left\lceil \frac{dst}{stp} \right\rceil + 1$ % Υπολογισμός αριθμού σημείων στην τομή

$xprf(i) = x1 + (i - 1) \cdot stp \cdot \cos(ang), i = 1, \dots, n$ % Υπολογισμός των πραγματικών συντεταγμένων κάθε σημείου

$yprf(i) = y1 + (i - 1) \cdot stp \cdot \sin(ang)$

for $i = 1, \dots, n$ % Για κάθε σημείο στη τομή

$addst = 0, adz = 0, adg = 0$ % Μηδενισμός των αθροιστών

for $j = 1, \dots, n$ % Για κάθε μετρημένο σημείο

$dd = \sqrt{(xb(j) - xprf(i))^2 + (yb(j) - yprf(i))^2}$ % Απόσταση ανάμεσα στο σημείο της τομής και του μετρημένου σημείου

if $dd \leq rint$ then % Εάν η απόσταση είναι μικρότερη της ακτίνας αναζήτησης

$addst = addst + \frac{1}{dd^2}, adz = adz + \frac{zb(j)}{dd^2}, adg = adg + \frac{gb(j)}{dd^2}$

end for

if $addst > 0$ then % Εάν έχει βρεθεί κάποιο σημείο μέσα στην ακτίνα αναζήτησης
% Υπολογίζονται τα ζητούμενα

$zp(i) = \frac{adz}{addst}, gp(i) = \frac{adg}{addst}, xp(i) = (i - 1) \cdot stp$

else

display " Increase rint" % Χρειάζεστε μεγαλύτερη ακτίνα σύγκλισης

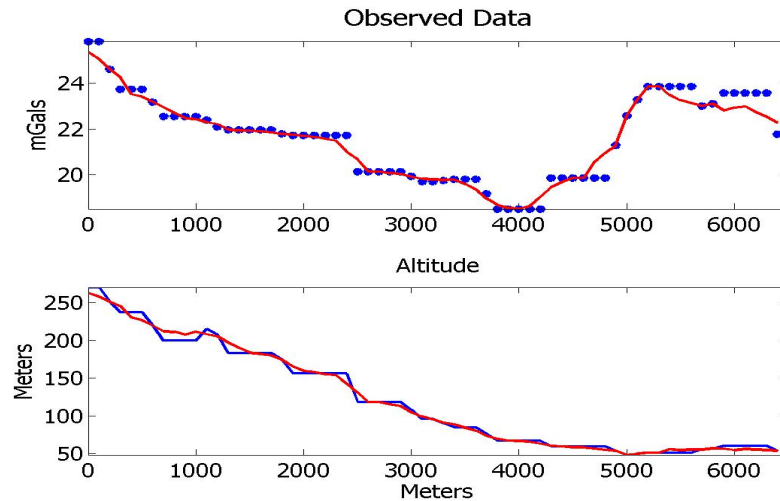
end if

end for

Όπως φαίνεται στο παρακάτω σχήμα, οι καμπύλες οι οποίες προκύπτουν από την εκτέλεση του αλγορίθμου με ακτίνα σύγκλισης 500 μέτρα (κόκκινο χρώμα) είναι πιο "ομαλές" από τις καμπύλες που προκύπτουν ζητώντας ακτίνα σύγκλισης 200 μέτρα (μπλε χρώμα).

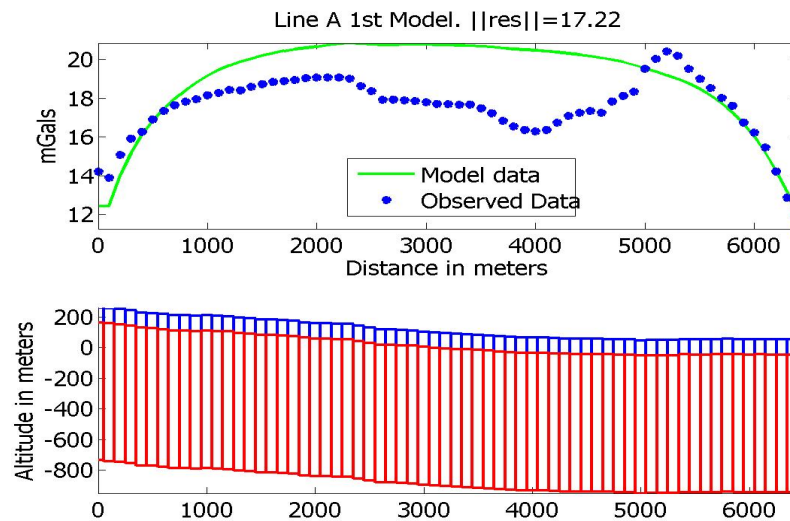
Οι εξομάλυνση των πρώτων οφείλεται στην μεγαλύτερη επιρροή του περιβάλλοντος. Αν μετράγαμε ακριβώς πάνω στα σημεία x_1, x_2, \dots, x_N θα είχαμε ακόμα πιο απότομες καμπύλες. Η εξομάλυνση βοηθάει στην ερμηνεία, οι αλγόριθμοι δεν πρέπει να προσεγγίζουν απότομες κλήσεις.

Στην περίπτωση μας θεώρησα σωστή μια ακτίνα σύγκλισης 500 μέτρων.



2.6 Μέθοδος Talwani.

Ως αρχικό μοντέλο επιλέγουμε ένα επιφανειακό στρώμα με πυκνότητα 2.2, με βάθος 100 μέτρα και από εκεί και κάτω ένα στρώμα με πυκνότητα 2.7:



όπου με μπλε χρώμα συμβολίζονται τα πρίσματα με πυκνότητα 2.2 και με κόκκινο αυτά με πυκνότητα 2.7. Ως εκτίμηση σφάλματος χρησιμοποιείται η Ευκλείδεια νόρμα των υπολοίπων.

Χρησιμοποιούμε μια επαναληπτική μέθοδο για την εκτίμηση του βάθους της επαφής μεταξύ δυο στρωμάτων με διαφορετική πυκνότητα :

Στον τύπο της διόρθωσης Bouguer 2.18 , η ποσότητα Δ_{3g} έχει διαστάσεις βαρύτητας. Σε κάθε σημείο μέτρησης υποθέτουμε ότι το **residual** = $41.91 \times 10^{-5} \cdot \text{πυκνότητα} \cdot \text{διαφορά ύψους}$. Έτσι, σε κάθε σημείο μέτρησης υπολογίζεται το καινούργιο βάθος του επιφανειακού στρώματος από τον τύπο :

$$\text{μεταβολή βάθους} = \frac{\text{residual}}{41.91 \times 10^{-5} \times \text{density}} \quad (2.20)$$

Για να πάμε στον αλγόριθμο της μεθόδου, χρειαζόμαστε την διαδικασία
Δημιουργία Αρχικού Μοντέλου (ΔΑΜ)

Για κάθε σημείο $i = 1, \dots, n$ παίρνουμε τις συντεταγμένες $xbox1_i = xp_i - \frac{stp}{2}$,
 $xbox2_i = xp_i + \frac{stp}{2}$, $zbox1_i = zp_i$, $zbox2_i = zp_i - 100$, $zbox3_i = zp_i - 1000$.

Οι συντεταγμένες $xbox1, xbox2, zbox1, zbox2$ δημιουργούν τα πρίσματα του επιφανειακού στρώματος ενώ οι $xbox1, xbox2, zbox2, zbox3$ δημιουργούν τα πρίσματα του υποβάθρου. Σε κάθε επανάληψη ουσιαστικά εκτιμάτε η μεταβλητή $zbox2$ η οποία δημιουργεί την επαφή μεταξύ των δύο στρωμάτων. Έχουμε επιλέξει μέγιστο βάθος ερμηνείας τα 1000 μέτρα. Το όργανο με το οποίο μετρήσαμε δεν μπορεί να λάβει σήμα από πιο βαθιά, το σήμα θα έχει μήκος κύματος ίδιο με το θόρυβο (σε όλες τις μετρήσεις, ακόμα και τις πειραματικές υπάρχει θόρυβος ο οποίος οφείλεται στην ακρίβεια των οργάνων, μικρή απόκλιση από τις βέλτιστες ρυθμίσεις, κτλ.). Γι' αυτό, το σήμα από βάθος μεγαλύτερο των 1000 μέτρων το εντάσσουμε στον θόρυβο.

Αλγόριθμος InvTalwani

Input: $xp_i, zp_i, gp_i, i = 1, \dots, n$ % Τα αποτελέσματα του αλγορίθμου InversDistance
 stp % Βήμα
 tol % Επιθυμητό Σφάλμα

Output: Σε κάθε επανάληψη παίρνουμε ως αποτέλεσμα το βάθος της επαφής των δυο στρωμάτων και την εκτίμηση του σφάλματος

Begin

ΔΑΜ

Διαδικασία Υπολογισμός ανταπόκρισης Μοντέλου (ΔΥΜ)

% Υπολογισμός του πίνακα A με τον τύπο του Banerjee
 % Για την ενότητα 2.3: $x_1 = xp - xbox1$, $x_2 = xp - xbox2$, $y_1 = zp - zbox1$, $y_2 = zp - zbox2$
 $A1 = \text{fCalcA}(xp, zp, xbox1, xbox2, zbox1, zbox2)$
 % Υπολογισμός βαρύτητας για το επιφανειακό στρώμα
 $g1 = A1 \times [0.2 \ 0.2 \ \dots \ 0.2]^T$
 % Ίδια διαδικασία για το στρώμα του υποβάθρου
 $A2 = \text{fCalcA}(xp, zp, xbox1, xbox2, zbox2, zbox3)$
 $g2 = A2 \times [0.7 \ 0.7 \ \dots \ 0.7]^T$
 % Η ανταπόκριση του μοντέλου είναι το άθροισμα των δύο διανυσμάτων
 $gc = g1 + g2$
 $err = ||gp - gc||$

Τέλος διαδικασίας

while $err \geq tol$
 for $i = 1, \dots, n$

$$res = gp_i - gc_i$$

$$dh = \frac{res}{0.0419 * 0.7}$$

$$zb = zbox2_i + dh$$

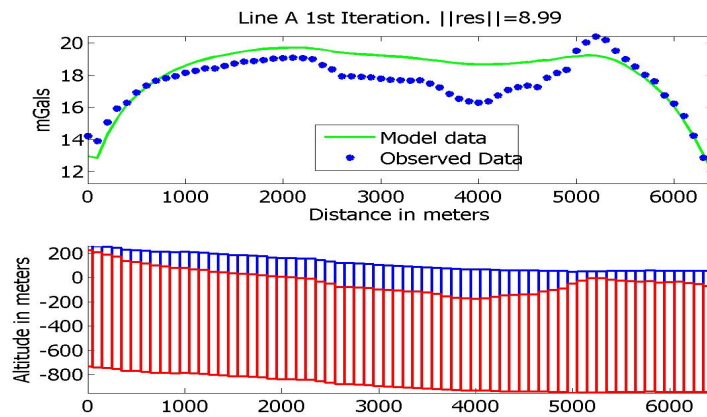
if $zb \geq zbox3_i$ and $zb \leq zbox1_i$ then $zbox2_i = zb$

end for

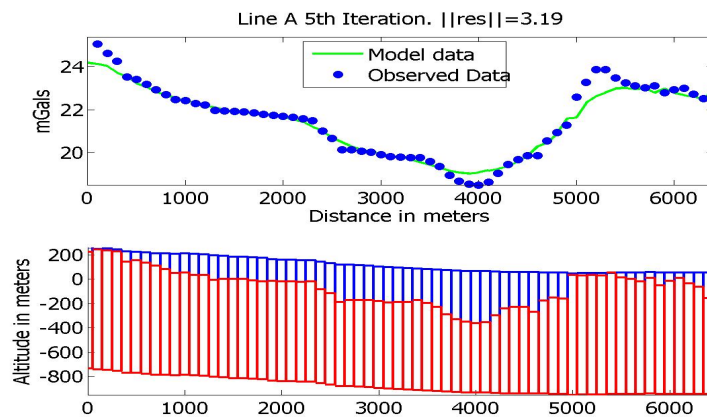
ΔΥΜ

end while

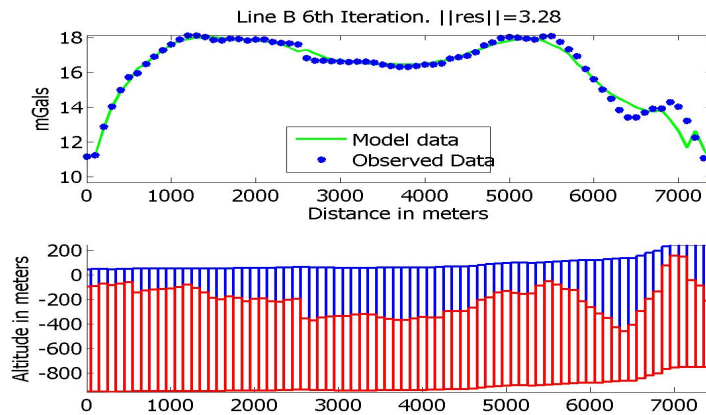
Το πρόγραμμα **plotTalw** δημιουργεί τη γραφική παράσταση.
Μετά την πρώτη επανάληψη θα έχουμε:



ενώ μετά από 5 επαναλήψεις:



Για την γραμμή **B** μετά από 6 επαναλήψεις:



Οι γραμμές A και B συναντιούνται σε ένα σημείο το οποίο αντιστοιχεί σε περίπου 3800 μέτρα από την αρχή της γραμμής A και 4400 μέτρα από την αρχή της γραμμής B. Σε αυτά τα σημεία, το πάχος του επιφανειακού στρώματος είναι περίπου 400 μέτρα. Οποιαδήποτε γραμμή τέμνει την γραμμή A, στο σημείο τομής τους οι δύο ερμηνείες θα έχουν το ίδιο βάθος.

2.7 Ανάγκη χρήσης μεθόδων Αριθμητικής Ανάλυσης

Μεγάλο μειονέκτημα της προηγούμενης μεθόδου είναι ότι εκτιμάει πολύ καλά μόνο την επαφή μεταξύ δύο στρωμάτων. Αν το μοντέλο μας ερμηνεύεται με περισσότερες τιμές πυκνότητας, όπως κάποιο έγκοιλο (κενό) μέσα σε κάποιο στρώμα, πρέπει να προσεγγίσουμε το πρόβλημα με διαφορετικό τρόπο.

2.7.1 Επιλογή Μοντέλου

Επιλέγω πάνω στη γραμμή των μετρήσεων ένα μοντέλο, αποτελούμενο από 13 σειρές πρισμάτων, οι 10 πρώτες σειρές με πάχος 50 μέτρα, η 11η σειρά με πάχος 100 μέτρα και οι δύο πιο βαθιές με πάχος 200 μέτρα. Προσπαθώ με αυτόν τον τρόπο να ερμηνεύσω τα 1000 μέτρα βάθους.

Έχω κάνει δοκιμές και με άλλα μοντέλα, πήρα αρκετά καλά αποτελέσματα, αρκεί το βάθος του μοντέλου να φτάνει στα 1000 μέτρα. Σε ένα μοντέλο με 7 σειρές πρισμάτων, με βάθος των επιφανειακών σειρών 100 μέτρα, δεν είχα τόσο καλή σύγκλιση. Είδα ότι αν το πάχος των αντίστοιχων πρισμάτων είναι 50 μέτρα, λόγω μεγαλύτερης διακριτικής ικανότητας, το μοντέλο προσεγγίζει καλύτερα τις διορθωμένες τιμές βαρύτητας. Επίσης, ήδη έχω 300 πρίσματα και ο πίνακας A θα έχει διάσταση 300×300 . Για καλύτερη διακριτική ικανότητα (π.χ. πάχος πρισμάτων 25 μέτρα) θα έχουμε ακόμα καλύτερη σύγκλιση, αλλά θα έχουμε επίσης και 1000 πρίσματα και μεγάλη σπατάλη υπολογιστικού χρόνου με τα αντίστοιχα σφάλματα υπολογισμών.

Όπως είδαμε στην εισαγωγή, οι τιμές της βαρύτητας στα σημεία $x_i, i = 1, \dots, n$ υπολογίζονται από το σύστημα γραμμικών εξισώσεων $g = Ax$ όπου $g \in R^n$ είναι οι τιμές βαρύτητας, $x \in R^n$ είναι οι προς εκτίμηση πυκνότητες και ο πίνακας $A \in R^{n \times n}$ είναι ένας γεωμετρικός παράγοντας ο οποίος υπολογίζεται από την απόσταση μεταξύ του σημείου μέτρησης x_i και κάθε πρισματος x_j , βρίσκοντας τα A_{ij} (ενότητα 2.3), άρα για διαφορετικό μοντέλο πρισμάτων έχουμε διαφορετικό πίνακα A . Ο τετραγωνικός πίνακας A έχει πολύ κακή κατάσταση. Για το μοντέλο μας ο πίνακας A έχει δείκτη κατάστασης $\kappa(A) := \|A\| \cdot \|A^{-1}\| = 8.4 \times 10^{11}$. Αυτό είναι ένα δύσκολο αριθμητικό πρόβλημα, γι' αυτό χρειάζεται να χρησιμοποιήσουμε μεθόδους αριθμητικής

ανάλυσης.

Έχουμε να λύσουμε το

Το πρόβλημα Αριθμητικής Ανάλυσης

$$\begin{aligned} \min \|Ax - g\| \\ A \in R^{n \times n}, g \in R^n - \text{γνωστά} \\ x \in R^n - \text{άγνωστα} \end{aligned} \quad (2.21)$$

με την προϋπόθεση $l_i \leq x_i \leq u_i, i = 1, \dots, n$

3 Εκτίμηση σφάλματος

3.1 Εισαγωγή

Έχουμε ένα σύστημα n γραμμικών εξισώσεων με n αγνώστους το οποίο γράφεται στην μορφή $Ax = g$. Θεωρούμε το x^* μια οποιαδήποτε λύση του συστήματος. Το πρόβλημα είναι να εξετάσουμε τότε το x^* είναι μια καλή προσέγγιση. Συνήθως η ποιότητα της προσέγγισης κρίνεται από τη νόρμα του διανύσματος των υπολοίπων (residual vector) $r = g - Ax^*$. Εύκολα διαπιστώνουμε ότι το σφάλμα $e = x - x^*$ σχετίζεται με το υπόλοιπο μέσω της σχέσης $Ae = r$ άρα δεν είναι εφικτό να υπολογίσουμε το σφάλμα από τα υπόλοιπα.

Παίρνουμε τις νόρμες στη τελευταία σχέση, $\|Ae\| = \|r\| \Rightarrow \|A\| \cdot \|e\| \geq \|r\| \Rightarrow \|e\| \geq \frac{\|r\|}{\|A\|}$ και στη $\|e\| = \|A^{-1}r\| \Rightarrow \|e\| \leq \|A^{-1}\| \cdot \|r\|$ και έχουμε τα όρια του σφάλματος

$$\frac{\|r\|}{\|A\|} \leq \|e\| \leq \|A^{-1}\| \cdot \|r\| \quad (3.1)$$

Όμως, αυτές οι εκτιμήσεις απαιτούν να γνωρίζουμε τη νόρμα του A ή του αντιστρόφου της. Για την Ευκλείδεια νόρμα, ο υπολογισμός των $\|A\|, \|A^{-1}\|$ είναι απαγορευτικός, η πολυπλοκότητα υπερβαίνει το n^3 .

Οι ανισότητες 3.1 μας δείχνουν ότι η $\|r\|$ μπορεί να χρησιμοποιηθεί ως μια καλή εκτίμηση του $\|e\|$ εάν οι $\|A\|$ και $\|A^{-1}\|$ είναι κοντά στο 1. Είδαμε όμως ότι ο πίνακας A της εφαρμογής έχει πολύ κακή κατάσταση, άρα η $\|r\|$ δεν μπορεί να χρησιμοποιηθεί ως εκτίμηση της $\|e\|$.

Ο Auchmuty [7] έδωσε μια εκτίμηση για όλες τις νόρμες στον τύπο του σφάλματος. Απέδειξε ότι η ποσότητα

$\tilde{\epsilon}_3 = \frac{\|r\|^2}{\|A^T r\|}$ είναι μια καλή προσέγγιση της Ευκλείδεια νόρμας του σφάλματος $\|e\|$. Όμως αυτή η εκτίμηση είναι ένα κάτω όριο για το $\|e\|$ διότι από την ανισότητα των Cauchy - Schwartz προκύπτει $(r, r) = (A^T r, x - x^*) \leq \|A^T r\| \cdot \|e\| \Rightarrow \|e\| \geq \frac{\|r\|^2}{\|A^T r\|}$.

3.2 Εκτίμηση του σφάλματος κατά Brezinski

Ο Brezinski στις εργασίες του [2, 3, 4] προτείνει διαφορετικές εκτιμήσεις για τη νόρμα του σφάλματος $e = x - x^*$ βασισμένες στη σχέση $r = Ae$. Αυτές οι εκτιμήσεις είναι έγκυρες για οποιοδήποτε μη ιδιάζων πίνακα A και

για μια προσέγγιση x^* αποτέλεσμα οποιασδήποτε μεθόδου, άμεσης ή επαναληπτικής.
 Παρουσιάζω το σκεπτικό του σε αυτή την ενότητα.

Θέτουμε

$$\begin{aligned} c_{2i} &= (r, (A^T A)^i r), & i \geq 0 \\ c_{2i} &= (r, (A A^T)^i r), & i \leq 0 \\ c_{2i+1} &= (r, A(A^T A)^i r), & i \geq 0 \\ c_{2i+1} &= (r, A(A A^T)^i r), & i \leq 0 \end{aligned}$$

Θα έχουμε

$$\begin{aligned} c_{-1} &= (r, A^{-1}r) = (r, e) = (Ae, e) \\ c_{-2} &= (A^{-1}r, A^{-1}r) = (e, e) \end{aligned}$$

Έτσι, το πρόβλημα για να βρούμε τα (r, e) και (e, e) μεταφέρεται στο να εκτιμήσουμε τα c_{-1}, c_{-2} . Στην εργασία [4] η εκτίμηση των c_{-1}, c_{-2} γίνεται μέσω των c_0, c_1, c_2 . Η προσέγγισή μας βασίζεται στην επέκταση (extrapolation) αυτών των τριών όρων της σειράς $(c_i), i = 0, 1, \dots$ στα σημεία c_0, c_1 και c_2 . Μπορούμε να χρησιμοποιήσουμε οποιαδήποτε μέθοδο επέκτασης για να εκτιμήσουμε αυτές τις δύο ποσότητες, αλλά η εκτίμηση είναι καλύτερη εάν η παρεμβάλουσα συνάρτηση (interpolating function) είναι κοντά στη ακριβή συμπεριφορά των c_i . Για να αντλήσουμε τις εκτιμήσεις των norm του σφάλματος, πρέπει πρώτα να αναλύσουμε τη συμπεριφορά των c_i , να δημιουργήσουμε μια διαδικασία επέκτασης (extrapolation procedure) καλά προσαρμοσμένη (well adapted) στο πρόβλημα.

Ας θεωρήσουμε τη "Ιδιόμορφη Παραγοντοποίηση Τιμών" (Singular Value Decomposition) του πίνακα A ,

$$A = USV^T$$

όπου $UU^T = VV^T = I$ και $S = \text{diag}(s_1, \dots, s_n)$ με $s_1 \geq s_2 \geq \dots \geq s_n > 0$ ιδιάζουσες τιμές του πίνακα A . Αν πάρουμε y ένα τυχαίο διάνυσμα και u_1, \dots, u_n τις στήλες του U , v_1, \dots, v_n τις στήλες του V , θα έχουμε

$$\begin{aligned} Ay &= \sum_{i=1}^n s_i(v_i, y)u_i \\ A^{-1}y &= \sum_{i=1}^n s_i^{-1}(u_i, y)v_i \end{aligned}$$

Εύκολα συμπεραίνουμε

$$c_0 = (r, r) = (U^T r, U^T r) = \sum_{i=1}^n (u_i, r)^2 \quad (3.2)$$

$$= (V^T r, V^T r) = \sum_{i=1}^n (v_i, r)^2 \quad (3.3)$$

$$c_1 = (r, Ar) = \sum_{i=1}^n s_i (u_i, r) (v_i, r) \quad (3.4)$$

$$c_2 = (Ar, Ar) = \sum_{i=1}^n s_i^2 (v_i, r)^2 \quad (3.5)$$

$$c_{-1} = (r, A^{-1}r) = \sum_{i=1}^n s_i^{-1} (u_i, r) (v_i, r) \quad (3.6)$$

$$c_{-2} = (A^{-1}r, A^{-1}r) = \sum_{i=1}^n s_i^{-2} (u_i, r)^2 \quad (3.7)$$

Έτσι, η νόρμα του σφάλματος μπορεί να υπολογιστεί από την 3.7 ($c_{-2} = (A^{-1}r, A^{-1}r) = (e, e)$) αλλά αυτός ο τύπος απαιτεί τη γνώση όλων των s_i, u_i που εμφανίζονται στο άθροισμα.

Προσεγγίσεις των c_{-1}, c_{-2} μπορούμε να πάρουμε κρατώντας μόνο τον πρώτο όρο του αθροίσματος της εξίσωσης 3.7. Επίσης, θα πάρουμε μόνο τον πρώτο όρο και στις εξισώσεις 3.2 -3.5. Θεωρώντας $\alpha = (u_1, r), \beta = (v_1, r), s = s_1$ πρέπει να βρούμε τα α, β, s τα οποία να ικανοποιούν τις συνθήκες παρεμβολής (interpolation conditions):

$$\begin{aligned} c_0 &= \alpha^2 = \beta^2 \\ c_1 &= \alpha\beta s \\ c_2 &= \beta^2 s^2 \end{aligned} \quad (3.8)$$

και μετά επεκτείνουμε για τις τιμές -1, -2 του δείκτη, άρα τα c_{-1}, c_{-2} προσεγγίζονται από:

$$\begin{aligned} c_{-1} &= \alpha\beta s^{-1} \\ c_{-2} &= \alpha^2 s^{-2} \end{aligned}$$

Επειδή το σύστημα 3.8 αποτελείται από 4 μη συμβατές εξισώσεις, πρέπει να μη λάβουμε υπόψη κάποιες από αυτές. Έτσι, το σύστημα 3.8 έχει μερικές λύσεις.

Λύνοντας το σύστημα

$$\begin{cases} c_0 = \beta^2 \\ c_1 = \alpha\beta s \\ c_2 = \beta^2 s^2 \end{cases}$$

παίρνουμε $\alpha = \frac{c_1}{\sqrt{c_2}}$, $s = \sqrt{\frac{c_2}{c_0}}$ καταλήγοντας στην εκτίμηση $c_{-2}^{(1)} = (e, e) = \frac{c_0 c_1^2}{c_2^2}$.

Από το σύστημα

$$\begin{cases} c_0 = \alpha^2 \\ c_0 = \beta^2 \\ c_2 = \beta^2 s^2 \end{cases}$$

παίρνουμε $\alpha = \sqrt{c_0}$, $s = \sqrt{\frac{c_2}{c_0}}$ καταλήγοντας στην εκτίμηση $c_{-2}^{(2)} = (e, e) = \frac{c_0^2}{c_2}$.

Επίσης, από το σύστημα

$$\begin{cases} c_0 = \alpha^2 \\ c_0 = \beta^2 \\ c_1 = \alpha \beta s \end{cases}$$

παίρνουμε $\alpha = \sqrt{c_0}$, $s = \frac{c_1}{c_0}$ καταλήγοντας στην εκτίμηση $c_{-2}^{(3)} = (e, e) = \frac{c_0^3}{c_1^2}$.

Από τις συνθήκες παρεμβολής 3.8 βλέπουμε ότι ισχύει $c_0 = \frac{c_1^2}{c_2}$, αντικαθιστώντας στον τύπο του $c_{-2}^{(1)}$ παίρνουμε

μια άλλη εκτίμηση, $c_{-2}^{(4)} = \frac{c_1^4}{c_2^3}$.

Τον τύπο του $c_{-2}^{(3)}$ τον πολλαπλασιάζουμε με c_0 και τον διαιρούμε με το ισοτιμό του $\frac{c_1^2}{c_2}$ για να πάρουμε

$$c_{-2}^{(5)} = \frac{c_0^3}{c_1^2} \cdot \frac{c_0}{\frac{c_1^2}{c_2}} = \frac{c_0^4 c_2}{c_1^4}.$$

Αυτές είναι οι εκτιμήσεις που έχει πάρει ο Brezinski στις εργασίες [2, 3, 4]. Επειδή αυτές οι εκτιμήσεις δεν είναι ανεξάρτητες, μπορούμε να πάρουμε περισσότερες συνδυάζοντας τις προηγούμενες.

Κατέληξε στη σειρά:

$$\begin{aligned} e_1^2 &= \frac{c_1^4}{c_2^3} \\ e_2^2 &= \frac{c_0 \cdot c_1^2}{c_2^2} & c_0 &= (r, r) \\ e_3^2 &= \frac{c_0^2}{c_2} & \text{όπου} & & c_1 &= (r, Ar) \\ e_4^2 &= \frac{c_0^3}{c_1^2} & & & c_2 &= (Ar, Ar) \\ e_5^2 &= \frac{c_0^4 \cdot c_2}{c_1^4} \end{aligned}$$

(3.9)

Όλοι οι τύποι των σφαλμάτων μπορούν να μαζευτούν στον συμπαγή τύπο :

$$e_i^2 = c_0^{i-1} \cdot (c_1^2)^{3-i} \cdot c_2^{i-4} \quad (3.10)$$

Ο συμπαγής αυτός τύπος γράφεται και ως:

$$e_v^2 = (c_0 c_1^{-2} c_2)^v \cdot (c_0^{-1} c_1^6 c_2^{-4}), \forall v \in \mathbb{R} \quad (3.11)$$

Επειδή $(r, Ar)(r, Ar) \leq (r, r)(Ar, Ar)$ συμπαιρόνουμε ότι $c_1^2 \leq c_0 c_2$ άρα $\rho = (c_0 c_1^{-2} c_2) \geq 1$ και η εξίσωση 3.11 γράφεται:

$$e_v^2 = \rho^v e_0^2 \quad (3.12)$$

η οποία είναι μια αύξουσα συνάρτηση του $v \in (-\infty, \infty)$. Συμπέρασμα, ισχύει:

$$e_1^2 \leq e_2^2 \leq e_3^2 \leq e_4^2 \leq e_5^2 \quad (3.13)$$

Αντικαθιστώντας τις παραμέτρους $c_i, i=1, 2, 3$ στον τύπο του e_3^2 και χρησιμοποιώντας την ανισότητα των Cauchy-Schwartz, παίρνουμε το αποτελέσματα :

$$e_3^2 = \frac{(r, r)^2}{(Ar, Ar)} = \frac{\|r\|^4}{\|Ar\|^2} \geq \frac{\|r\|^{4^2}}{\|A\|^2 \cdot \|r\|^2} = \frac{\|r\|^2}{\|A\|^2} \quad (3.14)$$

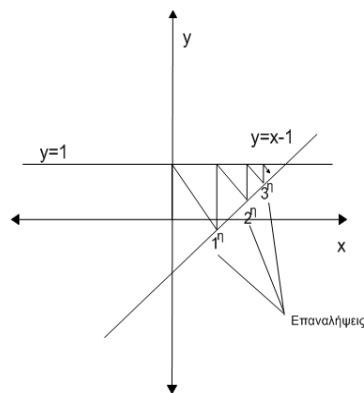
Στις δοκιμές μου χρησιμοποίησα των τύπο του $e_3^2 = \frac{c_0^2}{c_2}$. Για τον υπολογισμό του, χρειαζόμαστε n πράξεις για τον υπολογισμό του c_0 και n^2 για τον υπολογισμό του γινομένου Ar , δηλαδή έχουμε πολυπλοκότητα της τάξης του n^2 .

4 Επαναληπτικές μέθοδοι εκτίμησης παραμέτρων με υπό συνθήκη λύση

4.1 Αλγόριθμος Kaczmarz.

Προσαρμογή του για λύση του προβλήματος

Ο αλγόριθμος Kaczmarz [8] υλοποιείται εύκολα προς λύση του συστήματος $Ax = g$. Για να καταλάβουμε τον αλγόριθμο, σημειώνουμε ότι κάθε στήλη του συστήματος $Ax_i = g_i$ ορίζει ένα n -διάστατο επίπεδο στο R^n . Ο αλγόριθμος ξεκινάει με αρχική εκτίμηση $x^{(0)} = 0$ και προχωράει στην εκτίμηση $x^{(1)}$ μέσω της προβολής της αρχικής εκτίμησης $x^{(0)}$ πάνω στο επίπεδο που δημιουργεί η πρώτη στήλη του A . Στη συνέχεια προβάλετε πάνω στη δεύτερη στήλη του A , όταν προβάλετε και στην n -ωστή στήλη, αρχίζουμε πάλι από την πρώτη. Αυτός ο κύκλος επαναλαμβάνεται μέχρι να φτάσουμε στο επιθυμητό σφάλμα.



Παράδειγμα για το σύστημα

$$\begin{cases} y = 1 \\ x - y = 1 \end{cases}$$

του οποίου τη γραφική αναπαράσταση βλέπουμε παραπάνω, μετά από μερικές επαναλήψεις πλησιάζει την λύση η οποία είναι $(2, 1)$.

Για να υλοποιήσουμε τον αλγόριθμο χρειαζόμαστε έναν τύπο για τον υπολογισμό της προβολής ενός διανύσματος πάνω στο n -διάστατο επίπεδο που δημιουργείτε για κάθε $i = 1, \dots, n$ στήλη του A . Θεωρούμε A_i την i στήλη του A . Παίρνουμε το επίπεδο που ορίζεται από το $A_{i+1}x = g_{i+1}$. Επειδή το διάνυσμα A_{i+1}^T είναι κάθετο σε αυτό το επίπεδο, η ποσότητα η οποία ενημερώνει το $x^{(i)}$ από τους περιορισμούς της $i+1$ στήλης θα είναι ανάλογο του A_{i+1}^T ,

$$x^{(i+1)} = x^{(i)} + \beta A_{i+1}^T \quad (4.1)$$

Επειδή ισχύει $A_{i+1}x = g_{i+1}$, για να λύσουμε προς β παίρνουμε

$$A_{i+1}(x^{(i)} + \beta A_{i+1}^T) = g_{i+1} \quad (4.2)$$

Μετά από πράξεις καταλήγουμε

$$\beta = -\frac{A_{i+1}x^{(i)} - g_{i+1}}{A_{i+1}A_{i+1}^T} \quad (4.3)$$

άρα

$$x^{(i+1)} = x^{(i)} - \frac{A_{i+1}x^{(i)} - g_{i+1}}{A_{i+1}A_{i+1}^T} A_{i+1}^T \quad (4.4)$$

Δοκίμασα ως αρχικό μοντέλο όχι μηδενικές πυκνότητες αλλά κάποιο πιο κοντά στην πραγματικότητα. Αποτέλεσμα, το σφάλμα μικραίνει σε κάθε επανάληψη. Για να διατηρήσω τα όρια $l_i \leq x_i \leq u_i$, αν η καινούργια υπολογισμένη τιμή είναι εκτός ορίων, κρατάω την παλιά.

Στο αρχικό μοντέλο θεώρησα ότι τα πρίσματα σε βάθος μεγαλύτερο από 700 μέτρα έχουν πυκνότητα 2.50 - 2.70. Στο υποπρόγραμμα **pTakeBounds** υπάρχουν και οι υπόλοιποι περιορισμοί. Στην επιλογή των περιορισμών με βοήθησαν τα αποτελέσματα μιας άλλης γεωφυσικής μεθόδου¹ η οποία εφαρμόστηκε στην περιοχή της έρευνας.

Διαδικασία Input-Output (ΔΙΟ)

Input: $l_i, x_i, u_i, i = 1, \dots, n$ % Αρχικό μοντέλο
 $xp_i, zp_i, gp_i, i = 1, \dots, n$ % Αποτελέσματα Αλγορίθμου InversDistance
 $xbox1_i, xbox2_i, zbox1_i, zbox2_i, i = 1, \dots, n$ % Συντεταγμένες κάθε πρίσματος
Output: σε κάθε επανάληψη *iter* παίρνουμε τις εκτιμώμενες πυκνότητες $x_i^{(iter)}$
και υπολογίζεται το αντίστοιχο σφάλμα $ferr^{(iter)} = e_3^2$.
ReadData % Δημιουργία πρισμάτων. Αρχικό μοντέλο
pTakeBounds % Δημιουργία $l_i, u_i, i = 1, \dots, n$ κάτω, άνω ορίων

Αλγόριθμος InvKaczmarz

ΔΙΟ

Begin

$A = \text{fCalcA}(xp, zp, xbox1, xbox2, zbox1, zbox2)$

$k=0, ferr = 10^{10}$

while $ferr \geq tol$

$k = k + 1, i = \text{mod}(k, n + 1)$

$\beta = -\frac{A_i x^{(k)} - g_i}{A_i A_i^T} A_i^T$

for $j = 1, \dots, n$

if $x_j + \beta_j \geq l_j$ and $x_j + \beta_j \leq u_j$ then $x_j = x_j + \beta_j$

end for

$g_c = Ax$ % Υπολογισμός βαρύτητας για το τρέχων μοντέλο

$ferr = \text{fCalcErr}(g_p, g_c, A)$

end while

¹Ηλεκτρικές βυθοσκοπήσεις με τη διάταξη Schlumberger


```

function ferr=fCalcErr( $g_p, g_c, A$ )
 $r = g_p - g_c$  % Υπολογισμός υπολοίπων
 $d = Ar$ 
 $ferr = \frac{(rr^T)^2}{dd^T}$  % Υπολογισμός σφάλματος

```

Σε κάθε επανάληψη ο αλγόριθμος υπολογίζει το διάνυσμα $\beta = -\frac{A_{i+1}x^{(i)} - g_{i+1}}{A_{i+1}A_{i+1}^T}$. Για τον υπολογισμό του ο

H/Y χρησιμοποιεί $3n$ πράξεις, αποτέλεσμα πολύ ικανοποιητικό. Απλώς, χρειάζονται πολλές επαναλήψεις. Για τον υπολογισμό του σφάλματος, εκτελούνται n^2 πράξεις για τον υπολογισμό του $d = Ar$, οι υπόλοιπες πράξεις είναι τάξης n , μικρότερης του n^2 και δεν λαμβάνονται υπ' όψη. Συνολικά εκτελούνται n^2 πράξεις.

Υπάρχει μια ενδιαφέρουσα πρόταση των Thomas Strohmer και Roman Vershynin²[10] που ισχυρίζονται ότι σε κάθε επανάληψη δεν πρέπει να παίρνουμε σειριακά τις στήλες του πίνακα A αλλά τυχαία. Στην τυχαία επιλογή, κάθε στήλη συμμετέχει με πιθανότητα ανάλογη της Ευκλείδειας νόρμας της.

Αλγόριθμος InvKaczmarzRndz

```

Begin
ΔΙΟ
 $A = \text{fCalcA}(xp, zp, xbox1, xbox2, zbox1, zbox2)$ 
 $na_i = \|A_{i,:}\|, \quad i = 1, \dots, n$ 
 $sna = \sum_{i=1}^n na_i$ 
while  $ferr \geq tol$ 
     $r = \text{rand} \cdot sna$ 
     $ads = 0, \quad k = 0$ 
    while  $ads < r$ 
         $k = k + 1$ 
         $ads = ads + na_k$ 
    end while
     $i = k - 1$ 
     $\beta = -\frac{A_i x^{(k)} - g_i}{A_i A_i^T} A_i^T$ 
    for  $j = 1, \dots, n$ 
        if  $x_j + \beta_j \geq l_j$  and  $x_j + \beta_j \leq u_j$  then  $x_j = x_j + \beta_j$ 
    end for
     $g_c = Ax$ 
     $ferr = \text{fCalcErr}(g_p, g_c, A)$ 
end while

```

Στη εργασία [10] αποδεικνύεται ότι ο αλγόριθμος InvKaczmarzRndz συγκλίνει προς τη λύση x με μέσω σφάλμα:

$$E\|x_k - x\|_2^2 \leq \left(1 - \kappa(A)^2\right)^k \cdot \|x_0 - x\|_2^2 \quad (4.5)$$

²Department of Mathematics, University of California

Στην περίπτωση μας που έχουμε τους περιορισμούς $l_i \leq x_i \leq u_i$ η παραπάνω εκτίμηση δεν ισχύει.

4.2 Αλγόριθμος Levenberg-Marquardt

Η πρώτη εφαρμογή του αλγορίθμου των Levenberg-Marquardt έγινε για προβλήματα ελαχίστων τετραγώνων στη προσαρμογή καμπυλών: δίνοντας ένα σύνολο από ζευγάρια εμπειρικών δεδομένων, ανεξάρτητων και εξαρτημένων μεταβλητών (x_i, y_i) , $i=1, \dots, N$ (όπου N σύνολο ζευγαριών), γίνεται εκτίμηση των παραμέτρων β του μοντέλου της καμπύλης $f(x, \beta)$ έτσι ώστε το άθροισμα του τετραγώνου των διαφορών

$$S(\beta) = \sum_{i=1}^N [y_i - f(x_i, \beta)]^2 \quad (4.6)$$

να είναι ελάχιστο.

Όπως και άλλοι αλγόριθμοι ελαχιστοποίησης, ο αλγόριθμος των Levenberg-Marquardt είναι μια επαναληπτική διαδικασία. Ξεκινώντας από μια αρχική εκτίμηση β^0 σε κάθε επανάληψη φάχνουμε να βρούμε μια καινούργια εκτίμηση $\beta + \delta$ έτσι ώστε το σφάλμα της να είναι μικρότερο από τη προηγούμενη επανάληψη. Για την εκτίμηση του δ , οι συναρτήσεις $f(x_i, \beta + \delta)$, $i=1, \dots, N$ προσεγγίζονται από το

$$f(x_i, \beta + \delta) \approx f(x_i, \beta) + J_i \delta \quad (4.7)$$

όπου $J_i = \frac{\partial f(x_i, \beta)}{\partial \beta}$ είναι η κλήση της f με αναφορά στο β . Χρησιμοποιώντας τη 4.7 στην εξίσωση 4.6, παραγωγίζοντας και θέτοντας το αποτέλεσμα ίσον με μηδέν, παίρνουμε το αποτέλεσμα:

$$(J^T J) \delta = J^T (y - f(x, \beta)) \quad (4.8)$$

Η συνεισφορά του Levenberg ήταν στο να αντικαταστήσει αυτή την εξίσωση με μια "αποσβεννυμένη" έκδοση (damped version):

$$(J^T J + \lambda I) \delta = J^T (y - f(x, \beta)) \quad (4.9)$$

όπου η παράμετρος λ είναι θετική και εκτιμάται σε κάθε επανάληψη. Εάν η ελαχιστοποίηση του σφάλματος γίνεται με γρήγορο ρυθμό, για τη παράμετρο λ χρησιμοποιούμε μια μικρή τιμή και ο αλγόριθμος μοιάζει με τη μέθοδο Gauss-Newton. Εάν μια επανάληψη δεν μειώνει αρκετά το σφάλμα, δίνουμε μεγαλύτερες τιμές στη παράμετρο λ και ο αλγόριθμος πλησιάζει στη μέθοδο "ταχύτερης καθόδου" (steepest descent).

Ο αλγόριθμος του Levenberg έχει το μειονέκτημα ότι αν η τιμή της "αποσβεννυμένης" παραμέτρου λ είναι μεγάλη, η αντιστροφή του $(J^T J + \lambda I)$ δεν ενδείκνυται. Ο Marquardt σκέφτηκε ότι μπορούμε να προσαρμόσουμε κάθε παράμετρο της κλήσης σύμφωνα με τη καμπυλότητα, δίνοντας μεγαλύτερη μετατόπιση στη διεύθυνση που η κλήση είναι μικρότερη. Αντικατέστησε το μοναδιαίο πίνακα στον τύπο του Levenberg με τη διαγώνιο του Hessian πίνακα $J^T J$ με αποτέλεσμα τον αλγόριθμο τον Levenberg- Marquardt:

$$(J^T J + \lambda \text{diag}(J^T J)) \delta = J^T (y - f(x, \beta)) \quad (4.10)$$

4.3 Προσαρμογή του αλγορίθμου LM σε ModifiedLM

Είδαμε στην εισαγωγή ότι η ανταπόκριση του μοντέλου δίνεται από τον τύπο $g = Ax$ όπου $g \in R^n$ είναι οι διορθωμένες τιμές βαρύτητας, $x \in R^n$ είναι οι άγνωστες πυκνότητες των πρισμάτων (προς προσδιορισμό) και ο πίνακας $A \in R^{n \times n}$ είναι ένας γεωμετρικός παράγοντας ο οποίος υπολογίζεται από την απόσταση μεταξύ του σημείου μέτρησης και κάθε πρίσματος.

Στην περίπτωση μας όπου έχω $n=300$ ο πίνακας A έχει δείκτη κατάστασης $\kappa(A) = 8.4 \times 10^{11}$.

Το πρόβλημα το οποίο πρέπει να λύσουμε είναι το 2.21.

Στην εξίσωση 4.10 για την περίπτωση μας, η συνάρτηση $f(x) = Ax$ άρα ο πίνακας

$$J = \frac{\partial f(x)}{\partial x} = \frac{\partial (Ax)}{\partial x} = A$$

και παίρνουμε την διαδικασία:

$$x^{(k)} = x^{(k-1)} + \delta, \quad (A^T A + \lambda \text{diag}(A^T A))\delta = A^T (g - Ax^{(k-1)}) \quad (4.11)$$

όπου το διάνυσμα x_i^0 , $i = 1, \dots, n$ είναι το αρχικό μοντέλο και k ο αριθμός της επανάληψης. Επίσης, χρειαζόμαστε τα κάτω, άνω όρια l_i , u_i , $i=1, \dots, n$.

Παρατηρούμε ότι αν στη διαγώνιο του $A^T A$ προσθέσουμε μια σταθερά λ ο δείκτης κατάστασής του πέφτει. Στην περίπτωση μας, $\kappa(A^T A) = 8.4 \times 10^{11}$, $\kappa(A^T A + 0.015I) = 10^5$. Όσο η τιμή της παραμέτρου λ αυξάνεται ο δείκτης κατάστασης του $A^T A$ μειώνεται, π.χ. $\kappa(A^T A + 20I) = 100$.

Για να έχει το σύστημα 4.11 καλύτερη κατάσταση, χρησιμοποιούμε τη QR μορφή του πίνακα A , $A = QR$ όπου $Q \in R^{n \times n}$ είναι ορθομοναδιαίως πίνακας (ισχύει $QQ^T = I$) και ο $R \in R^{n \times n}$ είναι άνω τριγωνικός. Αντικαθιστώντας τα στην 4.11 και θέτοντας $g_n = Q^T g$ παίρνουμε την εξίσωση που χρησιμοποιούμε σε κάθε επανάληψη:

$$(R^T R + \lambda \text{diag}(R^T R))\delta = R^T (g_n - Rx^0). \quad (4.12)$$

Η λύση για το δ μπορεί να βρεθεί λύνοντας το σύστημα εξισώσεων 4.12 με τη μέθοδο του Gauss με μερική οδήγηση.

Είναι γνωστό ότι για τον υπολογισμό της QR μορφής του πίνακα A χρειάζονται $\frac{2n^3}{3}$ πράξεις,

αριθμός αρκετά μεγάλος, αλλά υπολογίζεται μόνο μια φορά, όχι σε κάθε επανάληψη. Επίσης, για τον υπολογισμό του $R^T R$, επειδή ο πίνακας R είναι άνω τριγωνικός, χρειάζονται $\frac{n^3}{2}$ πράξεις αλλά και πάλι, μια φορά μόνο.

Για να ισχύουν οι περιορισμοί του 2.21, στην επανάληψη k , θεωρούμε $w_i = x_i^{(k-1)} + \delta_i$. Αν ισχύει $l_i \leq w_i \leq u_i$, τότε μεταφέρουμε την τιμή της w_i στην $x_i^{(k)}$, διαφορετικά $x_i^{(k)} = x_i^{(k-1)}$. Έτσι προκύπτει το διάνυσμα $x^{(k)}$ με το αντίστοιχο σφάλμα.

Η επιλογή της παραμέτρου λ στην εξίσωση 4.12 είναι μια πρόκληση. Όπως είδαμε στη προηγούμενη παράγραφο (εξ. 4.9), η γενική στρατηγική είναι να ξεκινάμε από μια αρκετά μικρή τιμή και ανάλογα με την πορεία της κάθε επανάληψης, αυξάνεται η μειώνεται η τιμή της παραμέτρου.

Αλγόριθμος InvMLM

Begin

ΔΙΟ

$A = \text{fCalcA}(xp, zp, xbox1, xbox2, zbox1, zbox2)$

$[Q, R] = \text{qr}(A)$

$RtR_{ij} = \sum_{k=1}^j R_{ki}R_{kj}, \quad i, j=1, \dots, n$

$g_n = Q^T g_p, \quad g_c = Ax$

$ferr = \text{fCalcErr}(g_n, g_c, A)$

$\lambda = 0.015$ % Αρχική τιμή της παραμέτρου λ

$v = 1.5$ % Ρυθμός μεταβολής της παραμέτρου λ

while $ferr \geq tol$

$\lambda1 = \frac{\lambda}{v}, \quad \lambda2 = \frac{\lambda1}{v}$

$[ff1, gc1, x1] = \text{fCalcNewParam}(n, \lambda1, RtR, R, x, g_n, A, g)$

$[ff2, gc2, x2] = \text{fCalcNewParam}(n, \lambda2, RtR, R, x, g_n, A, g)$

if $ff1 < ferr$ or $ff2 < ferr$ then

$gc = gc1$

$ferr = ff1$

$\lambda = \lambda1$

if $ff2 < ff1$ then

$gc = gc2$

$ferr = ff2$

$\lambda = \lambda2$

else

while $ff1 > ferr$

$\lambda1 = \lambda1 \cdot v$

$[ff1, gc1, x1] = \text{fCalcNewParam}(n, \lambda1, RtR, R, x, g_n, A, g)$

end while

$gc = gc1$

$ferr = ff1$

$\lambda = \lambda1$

end if

end while

```

Function [ff, gc, xx]=fCalcNewParam (n, λ, RtR, R, x, gn, A, g)
Aq=RtR+diag(RtR)% Υπολογισμός αριστερού μέρους της εξίσωσης 4.12
bq=RT·(gn-Rx) % Υπολογισμός δεξιού μέρους της εξίσωσης 4.12
% Για τους παραπάνω υπολογισμούς χρειάζονται n+2n2 πράξεις
[d, ier]=fCalcGaus(Aq, bq)% Υπολογισμός του δ με τη μέθοδο Gaus με μερική οδήγηση
%Εάν η συνάρτηση επιστρέφει ier=1 η διαδικασία διακόπτεται
xx=x
for i=1,...,n
    wi=xi+di
    if wi ≥ li and wi ≤ ui then xxi=wi
end for
gc=A·xx
ff=fCalcErr(g, gc, A)

```

```

function[d, ier]=fCalcGaus(A, b)
ier=0
for k=1,...,n-1
% Εναλλάσσουμε τη στήλη k με αυτή που στη γραμμή k έχει τη μέγιστη τιμή.
% Επίσης εναλλάσσουμε και τα αντίστοιχα στοιχεία του διανύσματος b.
% Εάν για κάποιο k ισχύει A(k,k)=0 θέτουμε ier=1 και σταματάμε εδώ.
    for i=k+1,...,n
        if Akk ≠ 0
            p =  $\frac{A_{ik}}{A_{kk}}$ 
        else
            ier=1
            exit function
        for j=k+1,...,n
            Aij = Aij - p·Akj
        end for
        bi = bi - p·bk
    end for
end for

```

Οι πράξεις των συγκρίσεων και εναλλαγής δεδομένων δεν υπολογίζονται για τη πολυπλοκότητα ενός αλγορίθμου. Το υπόλοιπο κομμάτι της συνάρτησης χρειάζεται για την υλοποίησή του περίπου $\frac{n^3}{3}$ πράξεις.

Η συνάρτηση fCalcNewParam εκτελεί $2n^2 + n$ πράξεις, μικρότερη τάξη από τις πράξεις που χρειάζεται η συνάρτηση fCalcGaus, $\frac{n^3}{3}$ πράξεις. Σε κάθε επανάληψη, ανάλογα με τη κατάσταση που βρίσκετε, η συνάρτηση fCalcNewParam μπορεί να εκτελεστεί πολλές φορές (περισσότερο από 10 φορές). Λογικό είναι να πιστεύουμε ότι η πολυπλοκότητα θα είναι $\frac{10n^3}{3}$ πράξεις. Αλλά σε κάθε εκτέλεση της συνάρτησης fCalcNewParam ξεκινάμε τις πράξεις από την αρχή, χρησιμοποιούμε τους ίδιους πίνακες απλά με διαφορετική

τιμή της παραμέτρου λ . Μπορεί να χρειάζομαστε περισσότερο χρόνο αλλά η πολυπλοκότητα κάθε επανάληψης είναι $\frac{n^3}{3}$ πράξεις.

4.4 Μέθοδοι Κανονικοποίησης και Κανονικοποίηση Tichonov.

Για τις επόμενες 2 ενότητες εμπνεύστηκα από τις εργασίες [2, 3, 11].

Οι πρακτικές μέθοδοι για να λύσουμε το διακριτό πρόβλημα $Ax = g$ πρέπει να μειώσουν την επιρροή του θορύβου. Αυτές διαφέρουν μόνο στον τρόπο που προσδιορίζουν τις συναρτήσεις που φιλτράρουν τον θόρυβο. Ας θεωρήσουμε την SVD (singular value decomposition) παραγοντοποίηση,

$$A = \sum_{i=1}^n u_i \sigma_i v_i^T, \quad (4.13)$$

όπου τα u_i, v_i είναι τα αντίστοιχα αριστερά, δεξιά ιδιάζοντα διανύσματα, ορθογώνια, και οι ιδιάζουσες τιμές σ_i είναι μη αρνητικές, φθίνουσα σειρά έτσι ώστε $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_n \geq 0$. Το κοινό όλων των προβλημάτων που έχουν "κακή κατάσταση" (ill conditioned) είναι ότι ο πίνακας A έχει ένα σύμπλεγμα (cluster) από ιδιάζουσες τιμές κοντά στο μηδέν, σύμπλεγμα το οποίο μεγαλώνει με τη διάσταση του πίνακα.

Εύκολα διαπιστώνουμε ότι η λύση ελαχίστων τετραγώνων του μη περιορισμένου προβλήματος $Ax = g$ είναι

$$x_{LSQ} = \sum_{i=1}^n \frac{\alpha_i}{\sigma_i} v_i \quad \text{όπου} \quad \alpha_i = u_i^T g \quad (4.14)$$

Χρησιμοποιώντας ελάχιστα τετράγωνα έχουμε πρόβλημα διότι η λύσεις που θα αντιστοιχούν στις πολύ μικρές ιδιάζουσες τιμές μεγαλώνουν υπερβολικά την πληροφορία που περιέχουν σε σχέση με τις αντίστοιχες λύσεις των μεγαλύτερων ιδιάζουσών τιμών. Οποιαδήποτε πρακτική μέθοδο πρέπει να εισάγει συντελεστές f_i οι οποίοι φιλτράρουν τη λύση:

$$x_{filtered} = \sum_{i=1}^n f_i \frac{\alpha_i}{\sigma_i} v_i \quad (4.15)$$

Οι μέθοδοι κανονικοποίησης διαφέρουν μόνο στην επιλογή των συντελεστών f_i . Η πιο γνωστή μέθοδος κανονικοποίησης οφείλεται στον Tikhonov [12] ο οποίος επέλεξε τη λύση x_λ η οποία ελαχιστοποιεί το πρόβλημα

$$\min_x \{ \|Ax - g\|^2 + \lambda^2 \|x\|^2 \}. \quad (4.16)$$

Εδώ, η παράμετρος λ ελέγχει πόσο βάρος δίνουμε στην ελαχιστοποίηση της $\|x\|$ συγκριτικά με τη ελαχιστοποίηση της νόρμας των υπολοίπων (residual norm). Σε μερικές εφαρμογές δεν είναι κατάλληλη η ελαχιστοποίηση της Ευκλείδειας νόρμας αλλά προτιμότερο της ημινόρμας $\|Lx\|$, όπου L είναι μια διακριτή προσέγγιση κάποιας παραγώγου.

Μια άλλη μέθοδος κανονικοποίησης είναι η TruncatedSVD (TSVD) η οποία απλά "περικόβει" το άθροισμα του 4.14 σε ένα άνω όριο $k < n$, πριν αρχίσουν να κυριαρχούν οι μικρές ιδιάζουσες τιμές.

Ασφαλείς επαναληπτικές μέθοδοι για να λύσουν το πρόβλημα ελαχίστων τετραγώνων

$\min_x \|Ax - g\|$ έχουν ιδιότητες ελαχιστοποίησης.

Οι οικογένεια μεθόδων των συζυγών κλήσεων (Conjugate Gradient) ελαχιστοποιεί τη συνάρτηση ελαχίστων τετραγώνων πάνω σε μια επεκτινόμενη σειρά από υποχώρους

$$A_k = \text{span}\{A^T g, (A^T A)A^T g, \dots, (A^T A)^{k-1}A^T g\}$$

και κρατάει τη $\|x\|$ όσο μικρότερη γίνεται.

Μια άλλη δημοφιλή μέθοδο είναι της μέγιστης εντροπίας [13]. Βασίζεται στην ιδέα: αφού οι μετρήσεις g περιέχουν θόρυβο, δεν μπορούμε να ζητήσουμε από τις πυκνότητες x να τις αναπαράγουν ακριβώς, αλλά μόνο να τις προσεγγίσουμε σε κάποια όρια. Ανάμεσα σε όλα τα διάνυσματα x που ικανοποιούν τη συνθήκη

$$\|Ax - g\| \leq \lambda$$

η μέθοδος της μέγιστης εντροπίας επιλέγει το διάνυσμα του οποίου η εντροπία

$$s(x) = - \sum_{i=1}^n x_i \ln \left(\frac{x_i}{x_i^0} \right) - (x_i - x_i^0)$$

είναι μέγιστη. Το x_0 είναι μια αρχική εκτίμηση.

Στη συνέχεια θα ασχοληθούμε με τη μέθοδο κανονικοποίησης Tikhonov.

Είδαμε ότι εάν το γραμμικό σύστημα $Ax = g$ έχει κακή κατάσταση, η κανονικοποίηση Tikhonov προσπαθεί να βρει ένα διάνυσμα x_λ το οποίο να ελαχιστοποιεί το τετραγωνικό συναρτησιακό

$$J(\lambda, x) = \min_x \{ \|Ax - g\|^2 + \lambda^2 \|Lx\|^2 \} \quad (4.17)$$

Το διάνυσμα x_λ θα είναι η λύση του συστήματος

$$(C + \lambda^2 E)x_\lambda = A^T g \quad \text{όπου} \quad C = A^T A \quad \text{και} \quad E = L^T L \quad (4.18)$$

Επίσης, το x_λ είναι η λύση ελαχίστων τετραγώνων του συστήματος

$$\begin{bmatrix} A \\ \lambda L \end{bmatrix} x_\lambda = \begin{bmatrix} g \\ 0 \end{bmatrix} \quad (4.19)$$

Πολλαπλασιάζοντας και τα δύο μέρη με $[A^T, \lambda L^T]$ παίρνουμε την εξίσωση 4.18 .

Εάν θέσουμε $r_\lambda = g - Ax_\lambda$ μπορούμε να έχουμε:

$$A^T r_\lambda = \lambda^2 E x_\lambda \quad (4.20)$$

Επειδή από την εξίσωση 4.18 η λύση γράφεται ως

$$x_\lambda = (C + \lambda^2 E)^{-1} A^T g = (A^T A + \lambda^2 L^T L)^{-1} A^T g \quad (4.21)$$

είναι φανερό ότι

$$\lim_{\lambda \searrow 0} x_\lambda = x \quad \text{και} \quad \lim_{\lambda \rightarrow \infty} x_\lambda = 0.$$

Από την 4.13 έχουμε $A^T = \sum_{i=1}^n v_i \sigma_i u_i^T$ άρα για $L = I$ (κανονικοποίηση Tikhonov μηδενικού βαθμού) αντικαθιστώντας στην 4.21 έχουμε

$$x_\lambda = \left(\sum_{i=1}^n v_i \sigma_i u_i^T u_i \sigma_i v_i^T + \lambda^2 I \right)^{-1} \sum_{i=1}^n v_i \sigma_i u_i^T g \quad (4.22)$$

Τα διανύσματα u_i είναι ορθογώνια άρα $u_i^T u_i = I$. Επίσης, επειδή τα σ_i είναι οι ιδιάζουσες τιμές, θα έχουμε $v_i \sigma_i \sigma_i v_i^T = \sigma_i v_i v_i^T \sigma_i = \sigma_i^2$ και η εξίσωση 4.22 γίνεται

$$x_\lambda = \sum_{i=1}^n \frac{1}{\sigma_i^2 + \lambda_i^2} v_i \sigma_i u_i^T g \quad (4.23)$$

Πολλαπλασιάζοντας και διαιρώντας την 4.23 με σ_i καταλήγουμε στην

$$x_\lambda = \sum_{i=1}^n \frac{\sigma_i^2}{\sigma_i^2 + \lambda_i^2} \frac{u_i^T g}{\sigma_i} v_i \quad (4.24)$$

η οποία είναι η εξίσωση 4.15 με συντελεστές $f_i = \frac{\sigma_i^2}{\sigma_i^2 + \lambda_i^2}$.

Για κανονικοποίηση Tikhonov μεγαλύτερου βαθμού, χρησιμοποιούμε τη μέθοδο GSVD (Generalized SVD) [16] και ο πίνακας L της εξίσωσης 4.17 είναι διάφορος του μοναδιαίου πίνακα. Για κανονικοποίηση Tikhonov πρώτου βαθμού ο πίνακας L θα έχει τη μορφή

$$L = \begin{bmatrix} 1 & 0 & \cdots & 0 & 0 & 0 \\ -1 & 1 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & -1 & 1 & 0 \\ 0 & 0 & \cdots & 0 & 0 & 1 \end{bmatrix}$$

πίνακας ο οποίος παριστάνει τη διακριτή προσέγγιση της πρώτης παραγώγου. Δυστυχώς, ο ορισμός της GSVD με τους αντίστοιχους συμβολισμούς δεν έχει τυποποιηθεί επί του παρόντος. Η Matlab έχει ενσωματωμένη την συνάρτηση GSVD η οποία δέχεται ως είσοδο τους πίνακες A , L και επιστρέφει τους πίνακες U , V , X , C , S όπου U , V είναι ορθομοναδιαίοι πίνακες και οι C , S μη αρνητικοί διαγώνιοι πίνακες για τους οποίους ισχύουν:

$$\begin{aligned} A &= UCX^T \\ L &= VSX^T \\ C^T C + S^T S &= I \end{aligned}$$

Οι γενικευμένες ιδιάζουσες τιμές θα είναι

$$\gamma_i = \frac{c_{ii}}{s_{ii}} \quad \text{όπου} \quad 0 \leq \gamma_1 \leq \gamma_2 \leq \cdots \leq \gamma_n$$

Η GSVD λύση της εξίσωσης 4.17 θα είναι

$$x_\lambda = \sum_{i=1}^n \frac{\gamma_i^2}{\gamma_i^2 + \lambda_i^2} \frac{u_i^T g}{\gamma_i} v_i \quad (4.25)$$

η οποία είναι η εξίσωση 4.15 με συντελεστές $f_i = \frac{\gamma_i^2}{\gamma_i^2 + \lambda_i^2}$.

Για κανονικοποίηση Tikhonov δεύτερου βαθμού ο πίνακας που χρησιμοποιούμε είναι η διακριτή προσέγγιση της δεύτερης παραγώγου:

$$L = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 & 0 & 0 \\ -1 & 2 & -1 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & -1 & 2 & -1 \\ 0 & 0 & 0 & \cdots & 0 & 0 & 1 \end{bmatrix}$$

Εκτίμηση παραμέτρου λ

Μια μέθοδο για να εκτιμήσουμε τη παράμετρο κανονικοποίησης λ είναι η L-curve [14]. Η L-curve σχεδιάζει την καμπύλη των παραμέτρων $(\rho(\lambda), \eta(\lambda))$ όπου τα $\rho(\lambda) = \|x_\lambda\|$, $\eta(\lambda) = \|r_\lambda\|$ μετράνε τη διάσταση της κανονικοποιημένης λύσης και των αντίστοιχων υπολοίπων (residual) [15]. Η ιδέα είναι ότι μια καλή μέθοδο η οποία θα επιλέγει τη παράμετρο λ πρέπει να περιέχει πληροφορίες για τη διάσταση της λύσης όσο και να χρησιμοποιεί πληροφορίες για τη διάσταση των υπολοίπων. Αυτό βέβαια είναι φυσικό διότι αναζητούμε μια ισορροπία που θα κρατάει και τις δύο τιμές χαμηλές. Η L-curve έχει μια ευδιάκριτη μορφή που μοιάζει στο αγγλικό γράμμα L και η γωνία του εντοπίζεται ακριβώς εκεί που η λύση x_λ αλλάζει τη φύση της, από εκεί που υπερέρχουν τα σφάλματα κανονικοποίησης (π.χ. υπερεξομάλυνση) εκεί που υπερέρχουν τα σφάλματα των g . Έτσι, η γωνία της L-curve αντιστοιχεί σε μια καλή ισορροπία ανάμεσα στη ελαχιστοποίηση των διαστάσεων και η αντίστοιχη παράμετρος κανονικοποίησης λ είναι μια καλή επιλογή.

Πολλές ιδιότητες της L-curve για τη κανονικοποίηση Tikhonov εξετάζονται στην εργασία [14].

Έχει αποδειχτεί ότι κάτω από σίγουρες υποθέσεις η L-curve (ρ, η) έχει δύο χαρακτηριστικές πλευρές, μια "επίπεδη" πλευρά όπου στη λύση x_λ κυριαρχούν τα σφάλματα κανονικοποίησης και μια "κάθετη" πλευρά στην οποία τα x_λ κυριαρχούνται από τα σφάλματα μέτρησης.

Στα πειράματά μας βρήκαμε ότι τις περισσότερες φορές είναι πλεονέκτημα να εξετάζουμε τη L-curve σε λογαριθμική κλίμακα. Υπάρχει μια δυνατή δικαιολογία για αυτό. Αφού οι ιδιαίζουσες τιμές καλύπτουν μερικές τάξεις μεγεθών, η συμπεριφορά της L-curve φαίνεται καλύτερα σε λογαριθμική κλίμακα. Επίσης, η λογαριθμική κλίμακα τονίζει τις επίπεδες πλευρές της L-curve όπου οι μεταβολές των ρ ή η είναι μικρές σε σύγκριση με τις μεταβολές των άλλων μεταβλητών. Αυτές οι πλευρές της L-curve συχνά "στριμώχονται" κοντά στους άξονες σε γραμμική κλίμακα. Ένα ακόμα πλεονέκτημα της λογαριθμικής κλίμακας είναι ότι σε αλλαγές κλίμακας για

τις μετρήσεις g η L-curve απλά μετακινείται οριζόντια και κάθετα.

Το βιβλίο των R. Aster, B. Borchers, C. H. Thurber [17] περιέχει ένα CD με πιο πρόσφατη έκδοση των "Regularization tools: a Matlab package for analysis and solution of discrete ill-posed problems" του P. C. Hansen από την έκδοση που κυκλοφορεί στο Internet . Το CD περιέχει τη συνάρτηση **Tikhcstr** η οποία εκτελεί τη κανονικοποίηση Tikhonov με υπό συνθήκη λύση, ώστε να ισχύουν οι περιορισμοί $l_i \leq x_i \leq u_i$, $i = 1, \dots, n$. Με βοήθησε πολύ στο να εκτιμήσω σωστά τη κατάλληλη παράμετρο κανονικοποίησης λ εκτελώντας για διάφορες τιμές της παραμέτρου τη συνάρτηση tikhcstr ώστε να σχηματιστεί η L-curve.

Αν η γωνία δεν είναι τόσο ευδιάκριτη, δοκιμάζουμε και μια άλλη μέθοδο:

Παίρνουμε μια σειρά μ_i , $i = 1 \dots 5$ σε ίσες αποστάσεις στο διάστημα $[\lambda_{min}, \lambda_{max}]$. Για κάθε $\lambda_i = \mu_i$ υπολογίζουμε το διάνυσμα x_λ με τη συνάρτηση Tikhcstr και το αντίστοιχο σφάλμα e_3^2 . Το διάστημα που σχηματίζεται από το λ_i που δίνει το μικρότερο σφάλμα μαζί με το προηγούμενο και το επόμενο του, παίρνουν τη θέση των λ_{min} , λ_{max} και η διαδικασία επαναλαμβάνεται μέχρι το καινούργιο διάστημα να είναι αρκετά μικρό ή το σφάλμα των τελευταίων επαναλήψεων να συγκλίνει.

Για να εκτελέσει της συνάρτησης tikhcstr χρειαζόμαστε την

Διαδικασία Δημιουργίας Πινάκων ($\Delta\Delta\Pi$) της συνάρτησης tikhcstr

```
for  $i = 1, \dots, n$ 
     $L_{ii} = 1$ 
     $G_{ii} = 1$ ,  $G_{i+n,i} = -1$ 
     $d_{ii} = l_i$ ,  $d_{i+n,i} = -u_i$ 
end
```

Αλγόριθμος L_Er_Est

Input: $lmin$, $lmax$ % Εισάγονται τα όρια μέσα στα οποία να βρει το μικρότερο σφάλμα

Begin

ΔIO

$$x0 = \frac{l + u}{2}$$

$\Delta\Delta\Pi$

$nc = 5$

$tol = 10^{-10}$, $tolX = 10^{-10}$ % Όρια ανοχής για το διάστημα και τα σφάλματα

$iter = 0$, $iout = 0$

while $iout = 0$

$$stp = \frac{lmax - lmin}{nc - 1}$$

$$\mu_i = lmin + (i - 1) \cdot stp, \quad i = 1, \dots, nc$$

$iter = iter + 1$

$[x, rho, eta] = \text{tikhcstr}(A, g, G, d, L, \mu, x0)$

for $i = 1 : nc$

$$gc = A \cdot x_{:,i}$$

$$ferr_i = \text{fCalcErr}(g, gc, A)$$

```

end for
minerr = ferr1, ilc = 1
for j = 2, ..., n
    if ferrj < minerr then
        minerr = ferrj, ilc = j
    end for
if ilc > 1 then lmin = μilc-1 else lmin = μ1
if ilc < nc then lmax = μilc+1 else lmax = μnc
lliter = μilc, eeiter = minerr
if iter > 1 then
    if lliter - lliter-1 < tol then iout = 1
    if iter > 2 then
        if |eeiter-1 - eeiter| < tolX and |eeiter-2 - eeiter-1| < tolX then iout = 1
    end while

```

4.5 Τεχνικές παρεκβολής για γραμμικά συστήματα σε κακή κατάσταση

Στην εργασία [19] υπάρχουν τεχνικές αφαίρεσης της παραμέτρου λ από την λύση της κανονικοποίησης Tichonov, εξίσωση 4.25, $\bar{x}_\lambda = \sum_{i=1}^n \frac{\gamma_i^2}{\gamma_i^2 + \lambda_i^2} \frac{u_i^T g}{\gamma_i} \bar{v}_i$. Αυτή η λύση μπορεί να γραφτεί ως μια ρητή συνάρτηση:

$$R(\lambda) = \sum_{i=1}^n \frac{a_i}{b_i + \lambda} \bar{v}_i \quad (4.26)$$

όπου a_i, b_i (άγνωστοι) προς προσδιορισμό παράμετροι και τα \bar{v}_i είναι διανύσματα γραμμικά ανεξάρτητα, $i=1, \dots, n$. Οι $2n$ σε αριθμό άγνωστοι παράμετροι προσδιορίζονται υποθέτοντας ότι για κάποιες δοκιμές k θα ισχύει:

$$\bar{x}_k(\lambda) = R_k(\lambda) \quad (4.27)$$

Τότε μπορούμε να παρεκβάλουμε στο σημείο $\lambda=0$:

$$\bar{x}(0) = \sum_{i=1}^n \frac{a_i}{b_i} \bar{v}_i \quad (4.28)$$

Ας θεωρήσουμε δύο διαδοχικές δοκιμές, τις

$$\bar{x}_k = \sum_{i=1}^n \frac{a_i}{b_i + \lambda_k} \bar{v}_i, \quad \bar{x}_{k+1} = \sum_{i=1}^n \frac{a_i}{b_i + \lambda_{k+1}} \bar{v}_i \quad (4.29)$$

και τα διανύσματα $\bar{w}_i, i=1, \dots, n$ έτσι ώστε $(\bar{v}_i, \bar{w}_j) = \delta_{ij}$ (φυσικά, αφού τα v_i είναι ορθοκανονικά ισχύει $w_i = v_i$). Μετά πολλαπλασιάζουμε τις 4.29 με $\bar{w}_j, j=1, \dots, n$ για να πάρουμε :

$$(\bar{x}_k, \bar{w}_j) = \frac{a_j}{b_j + \lambda_k}, \quad (\bar{x}_{k+1}, \bar{w}_j) = \frac{a_j}{b_j + \lambda_{k+1}} \quad (4.30)$$

Από τις 4.30 προκύπτει

$$(b_j + \lambda_n)(\bar{x}_k, \bar{w}_j) = (b_j + \lambda_{k+1})(\bar{x}_{k+1}, \bar{w}_j) \quad (4.31)$$

για να καταλήξουμε

$$b_j = -\frac{\lambda_{k+1}(\bar{x}_{k+1}, \bar{w}_j) - \lambda_k(\bar{x}_k, \bar{w}_j)}{(\bar{x}_{k+1}, \bar{w}_j) - (\bar{x}_k, \bar{w}_j)}, \quad j=1, \dots, n. \quad (4.32)$$

Εύκολα βρίσκουμε και τα

$$a_j = (b_j + \lambda_k)(\bar{x}_k, \bar{w}_j), \quad j=1, \dots, n \quad (4.33)$$

άρα μπορούμε να βρούμε τη λύση 4.28 . Αυτή η λύση είναι ίδια με την λύση ελαχίστων τετραγώνων 4.14, μόνο που έχουμε περιορισμούς στη λύση .

4.6 Λύση του MLM με SVD

Όπως στη μέθοδο MLM, ξεκινώντας από ένα αρχικό μοντέλο x_0^i , $i=1, \dots, n$ και τα αντίστοιχα όριά τους $l_i \leq x_i \leq u_i$, πρέπει να βρούμε μια καινούργια εκτίμηση $x^0 + \delta$ έτσι ώστε το σφάλμα να ελαττώνεται σε κάθε επανάληψη. Λύνοντας την εξίσωση 4.11 με τη μέθοδο SVD (αντί $\text{diag}(A^T A)$ χρησιμοποιούμε τον μοναδιαίο I) έχουμε το αποτέλεσμα :

$$d_\lambda^{(k)} = \sum_{i=1}^n \frac{s_i}{s_i^2 + \lambda^2} u_i^T r v_i \quad (4.34)$$

όπου ο αριθμός k δηλώνει τον αριθμό της επανάληψης και η μεταβλητή r περιέχει τα residual.

Η υπόλοιπη διαδικασία είναι ίδια όπως στη μέθοδο MLM, με τη διαφορά ότι η παράμετρος λ υπολογίζεται σε κάθε επανάληψη με τη βοήθεια της συνάρτησης pGoldSec (θα τη δούμε στην επόμενη ενότητα).

Αλγόριθμος MLMSVD

Begin

ΔΙΟ

$A = \text{fCalcA}(xp, zp, xbox1, xbox2, zbox1, zbox2)$

$g_c = Ax, \quad fold = \text{fCalcErr}(g, g_c, A)$

$iter = 1, \quad ferr_1 = fold$

$iout = 1, \quad [u, s, v] = \text{svd}(A)$

$\lambda_a = 10^{-3}, \quad \lambda_c = 100$

while $iout$

$iter = iter + 1$

$r = g - g_c$

$xst = x, \quad \text{pGoldSec} \quad \% \text{Υπολογισμός της βέλτιστης παραμέτρου } \lambda$

$d = \sum_{i=1}^n \frac{s_i}{s_i^2 + \lambda^2} u_i^T r v_i \% \text{Υπολογισμός των ενημερώσεων για το διάνυσμα } x$

$w = x + d$

 for $i = 1 \dots, n$

 if $w_i \geq l_i$ and $w_i \leq u_i$ then $x_i = w_i$

 end for

$g_c = Ax, \quad ferr_{iter} = \text{fCalcErr}(g, g_c, A)$

 if $ferr_{iter} \leq tol$ then $iout = 0$

end while

Εκτίμηση παραμέτρου λ

Όπως είδαμε στην ενότητα 4.3, βάση της μεθόδου Levenberg–Marquardt ξεκινάμε από μια χαμηλή τιμή για τη παράμετρο λ , ας πούμε $\lambda_0 = 0.015$. Στο τέλος της επανάληψης, αν βρεθεί σφάλμα μικρότερο από της προηγούμενης επανάληψης, η παράμετρος λ διαιρείται με έναν αριθμό v ($v > 0$, ας είναι $v = 1.5$) διαφορετικά πολλαπλασιάζετε με τον ίδιο αριθμό.

Επέλεξα σε κάθε επανάληψη να γίνεται εκτίμηση της παραμέτρου με τον αλγόριθμο Golden Section Search (GSS) [9]. Τον δοκίμασα και σε παλαιότερες εφαρμογές και πήρα ικανοποιητικά αποτελέσματα.

Υποθέτω ότι η παράμετρος λ βρίσκεται στο διάστημα (λ_a, λ_c) . Επιλέγουμε ένα ενδιάμεσο σημείο λ_b . Θα έχω ελάχιστο στο διάστημα (λ_a, λ_c) εάν το σφάλμα στο σημείο λ_b είναι μικρότερο και από το σφάλμα στο σημείο λ_a και από του λ_c . Επιλέγουμε ένα σημείο λ_x μεταξύ λ_a, λ_b ή μεταξύ λ_b, λ_c . Εάν το σφάλμα στο σημείο λ_x είναι μεγαλύτερο από αυτό του λ_b επιλέγω την καινούργια τριάδα σημείων $\lambda_a, \lambda_b, \lambda_x$ διαφορετικά η τριάδα θα είναι $\lambda_b, \lambda_x, \lambda_c$. Και στις δύο περιπτώσεις, στο μεσαίο σημείο της τριάδας αντιστοιχεί το μικρότερο σφάλμα από τα σημεία που έχω εξετάσει. Επίσης, κάθε φορά το προς εξέταση διάστημα είναι μικρότερο από το προηγούμενο. Η διαδικασία επαναλαμβάνεται μέχρι το καινούργιο διάστημα να είναι αρκετά μικρό.

```

function  $\lambda = \text{Golden}(\lambda_a, \lambda_b, \lambda_c, tol)$ 
 $r = 0.61803399, \quad c = 1 - r$ 
 $\lambda_b = \frac{\lambda_a + \lambda_c}{2}$ 
 $x0 = \lambda_a, \quad x3 = \lambda_c$ 
if  $|\lambda_c - \lambda_b| > |\lambda_b - \lambda_a|$ 
     $x1 = \lambda_b, \quad x2 = \lambda_b + c \cdot (\lambda_c - \lambda_b)$ 
else
     $x2 = \lambda_b, \quad x1 = \lambda_b - c \cdot (\lambda_b - \lambda_a)$ 
 $f1 = \text{func}(x1), \quad f2 = \text{func}(x2)$ 
 $iout = 1$ 
while  $iout$ 
    if  $f2 < f1$  then
         $x0 = x1, \quad x1 = x2$ 
         $x2 = r \cdot x1 + c \cdot x3$ 
         $f0 = f1, \quad f1 = f2$ 
         $f2 = \text{func}(x2)$ 
    else
         $x3 = x2, \quad x2 = x1$ 
         $x1 = r \cdot x2 + c \cdot x0$ 
         $f3 = f2, \quad f2 = f1$ 
         $f1 = \text{func}(x1)$ 
    if  $|x3 - x0| \leq tol \cdot (x1 + x2)$  then  $iout = 0$ 
end while
if  $f1 < f2$  then  $\lambda = x1$  else  $\lambda = x2$ 

```

```

function  $f = \text{func}(\lambda)$ 

```

% Όλοι οι πίνακες που χρειάζονται στη συνάρτηση μεταφέρονται από το κύριο πρόγραμμα μέσω

% Global μεταβλητών

$$d = \sum_{i=1}^n \frac{s_i}{s_i^2 + \lambda^2} u_i^T r v_i$$

```

 $xx = x$ 

```

```

for  $i = 1, \dots, n$ 

```

$$w_i = x_i + d_i$$

```

    if  $l_i \leq w_i$  and  $w_i \leq u_i$  then  $xx_i = w_i$ 

```

```

end for

```

$$gg = A \cdot xx$$

$$f = \text{fCalcErr}(g, gg, A)$$

5 Αριθμητικά Αποτελέσματα

5.1 Εκτίμηση σφάλματος

Θέλουμε να δούμε αν όντως ο τύπος για τον υπολογισμό του σφάλματος $e_3^2 = \frac{(r, r)^2}{(Ar, Ar)}$ δίνει καλύτερα αποτελέσματα από τον τύπο υπολογισμού των υπολοίπων $r = \|g - Ax^*\|$. Στην εργασία [4] υπάρχει μια ενδιαφέρουσα πρόταση: θέλουμε να βρούμε τα όρια του $\frac{\|r\|}{\|e\|}$. Διαιρούμε τις ανισότητες 3.1 με $\|r\|$ και η αντιστροφός τους είναι:

$$\frac{1}{\|A^{-1}\|} \leq \frac{\|r\|}{\|e\|} \leq \|A\| \quad (5.1)$$

Προκύπτουν οι ακόλουθες περιπτώσεις σύμφωνα με τις σχετικές θέσεις των διαστημάτων

$\left[\frac{1}{\|A^{-1}\|}, \|A\| \right]$ και $\left[\frac{1}{\kappa}, \kappa \right]$ (όπου κ ο δείκτης κατάστασης του πίνακα A):

- (1) $\|A\| \leq \kappa^{-1}$: τα δύο διαστήματα είναι ανεξάρτητα. Αντικαθιστώντας στην 5.1 παίρνουμε $\|r\| \leq \frac{\|e\|}{\kappa}$ και επειδή $\kappa \geq 1$ συμπαιρένουμε $\|r\| < \|e\|$.
- (2) $\frac{1}{\|A^{-1}\|} \leq \kappa^{-1} \leq \|A\| \leq \kappa$: Αυτές οι συνθήκες είναι ισοδύναμες με $\|A\| \leq 1 \leq \|A^{-1}\|$ και $\|A\|^{-1} \leq \kappa$. Συνεπώς, $\|A\| \leq 1$ και $\|r\| < \|e\|$.
- (3) $\frac{1}{\|A^{-1}\|} \leq \kappa^{-1} \leq \kappa \leq \|A\|$: αυτές οι συνθήκες μας λένε ότι και οι δύο νόρμες, $\|A\|$ και $\|A^{-1}\|$ είναι μικρότερες ή ίσες του 1. Αυτό είναι δυνατόν μόνο αν και οι δύο νόρμες ισούνται με 1 και συμπαιρένουμε ότι η $\|r\|$ (εξίσωση 3.1) και η $e_3 = \sqrt{\|rr^T\|}$ ισούται με $\|e\|$.
- (4) $\kappa^{-1} \leq \frac{1}{\|A^{-1}\|} \leq \kappa \leq \|A\|$: αυτές οι ανισότητες είναι ισοδύναμες με $\|A^{-1}\| \leq 1 \leq \|A\|$. Συνεπώς (λόγω της 5.1) $\|e\| \leq \|A^{-1}\| \cdot \|r\|$ άρα $\|e\| \leq \|r\|$.
- (5) $\kappa \leq \frac{1}{\|A^{-1}\|}$: τα δύο διαστήματα είναι ανεξάρτητα. Επειδή $\frac{1}{\|A^{-1}\|} \geq \kappa \geq 1$ αντικαθιστώντας στην 5.1 παίρνουμε $\|e\| \leq \|r\|$.
- (6) $\kappa^{-1} \leq \frac{1}{\|A^{-1}\|} \leq \|A\| \leq \kappa$: αυτές οι ανισότητες μας οδηγούν στο συμπέρασμα ότι και οι δύο νόρμες, $\|A\|$ και $\|A^{-1}\|$ είναι μεγαλύτερες ή ίσες του 1. Αποτέλεσμα, και οι δύο ποσότητες, $\|r\|$ και e_3 μπορούν να προσεγγίζουν το $\|e\|$.

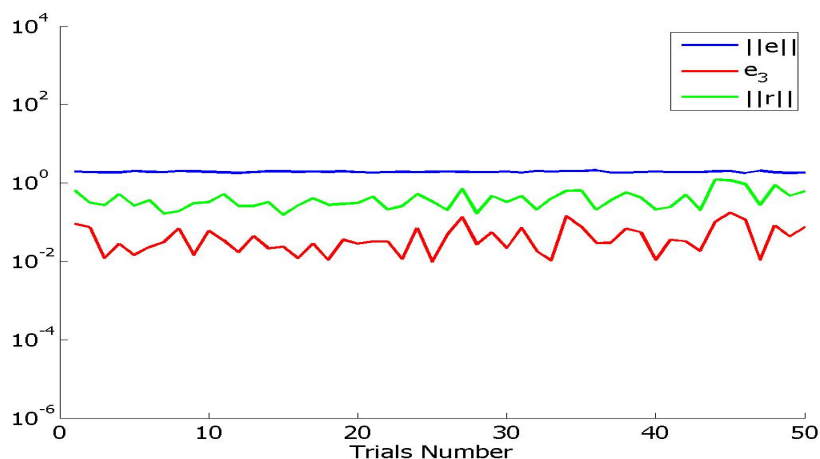
Διαδικασία 1

Για 50 επαναλήψεις, παίρνω ένα τυχαίο αρχικό μοντέλο μέσα στα όριά μου και υπολογίζω το $g = Ax$. Παίρνω ως αρχική εκτίμηση το μέσο όρο των ορίων και εκτελώ μια επανάληψη με τον αλγόριθμο MLM (ενότητα 4.3).

Αλγόριθμος CheckErr1

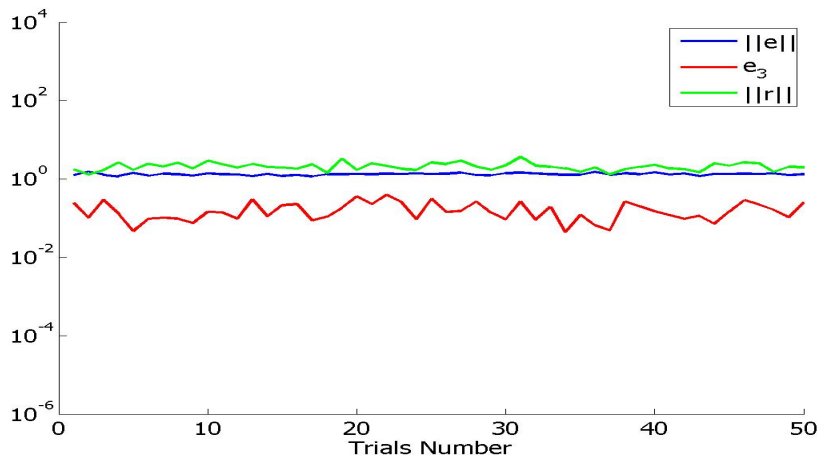
```
Begin
for  $i = 1, \dots, 50$ 
     $x = l + \text{rand} \cdot (u - l)$ 
     $g = Ax$ 
     $x^{(0)} = \frac{l+u}{2}$ 
     $x^{(1)} = \text{MLM}(x^{(0)})$ 
     $g_c = Ax^{(1)}$ 
     $nE_i = \|x - x^{(1)}\|$ 
     $nE3_i = \text{fCalcErr}(g, g_c, A)$ 
     $nR_i = \|g - g_c\|$ 
end for
```

Ξεκινάμε με τον πίνακα A_1 διάστασης 300×300 με δείκτη κατάστασης $\kappa = 8.4 \times 10^{11}$. Έχει $\|A_1\| = 50$, $\frac{1}{\|A_1^{-1}\|} = 6 \times 10^{-11}$ και $\kappa^{-1} = 1.2 \times 10^{-12}$. Το πρόβλημα ανήκει στην κατηγορία (6), άρα δεν μπορούμε να αποφασίσουμε εκ των προτέρων ποια ποσότητα από τις $\|r\|$, e_3 προσεγγίζει καλύτερα το σφάλμα $\|e\|$. Εκτελώντας τη διαδικασία 1 πήρα τα αποτελέσματα:



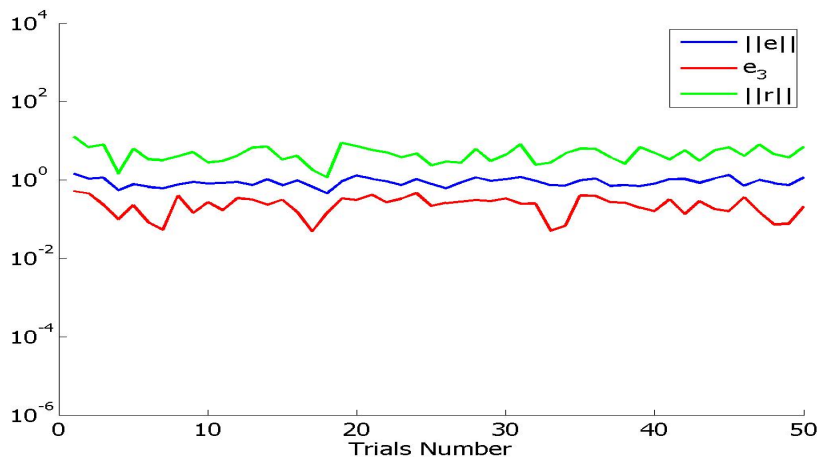
Παρατηρούμε ότι $e_3 < \|r\|$ και ότι το $\|r\|$ είναι πιο κοντά στο $\|e\|$ αλλά έχουν μεγάλη απόκλιση. Επίσης ισχύει $e_3 < \|r\| < \|e\|$.

Συνεχίζω με πίνακα A_2 διάστασης 150×150 (5 σειρές πρισμάτων) με δείκτη κατάστασης $\kappa = 5 \times 10^7$. Έχει $\|A_2\| = 69$, $\frac{1}{\|A_2^{-1}\|} = 1.4 \times 10^{-6}$ και $\kappa^{-1} = 2 \times 10^{-8}$. Επίσης ανήκει στην κατηγορία (6). Εκτελώντας τη διαδικασία 1 πήρα τα αποτελέσματα:



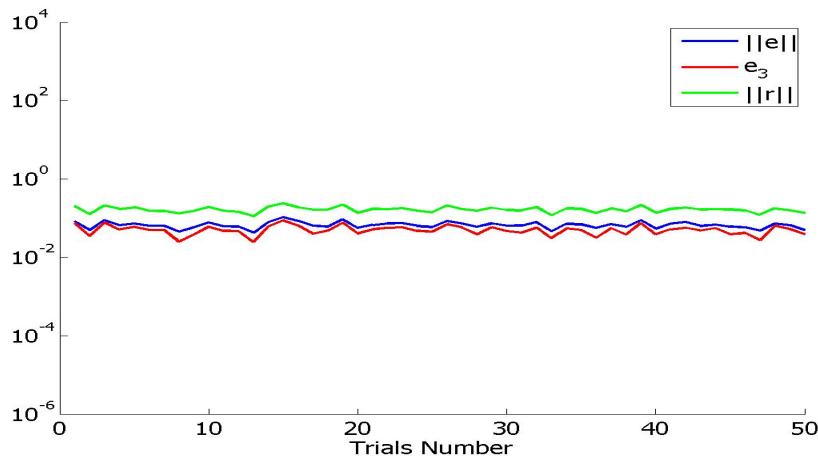
Παρατηρούμαι ότι και πάλι ισχύει $e_3 < \|r\|$ με τη διαφορά ότι το $\|r\|$ προσεγγίζει καλύτερα το $\|e\|$ από την εκτέλεση με τον πίνακα A_1 . Όμως η γραφική παράσταση είναι σε λογαριθμική κλίμακα και μας ξεγελάει. Αν ήταν σε γραμμική κλίμακα, θα βλέπαμε ότι τα $\|r\|$ και e_3 ισαπέχουν από το $\|e\|$. Το αποδεικνύει και ο πίνακας των σχετικών σφαλμάτων λίγο πιο κάτω.

Δοκιμάζω την ίδια διαδικασία με τον πίνακα A_3 διάστασης 100×100 (2 σειρές πρισμάτων) με δείκτη κατάστασης $\kappa = 2.5 \times 10^4$, $\|A_3\| = 112$, $\frac{1}{\|A_3^{-1}\|} = 3.8 \times 10^{-5}$ και $\kappa^{-1} = 4.3 \times 10^{-3}$. Και αυτός ο πίνακας ανήκει στη κατηγορία (6).



Παρατηρούμε ότι οι ποσότητες $\|r\|$ και e_3 έχουν την ίδια απόκλιση από το σφάλμα $\|e\|$ και ότι σε μερικές περιπτώσεις το e_3 προσεγγίζει καλύτερα το $\|e\|$.

Ένας πίνακας A_4 διάστασης 50×50 (1 σειρά πρισμάτων) με δείκτη κατάστασης $\kappa=20$, $\|A_4\| = 36$, $\frac{1}{\|A_4^{-1}\|} = 1.8$ και $\kappa^{-1} = 5 \times 10^{-2}$ ανήκει στη κατηγορία (4) και είμαστε σίγουροι ότι ισχύει $\|e\| \leq \|r\|$. Το βεβαιώνει και η γραφική του παράσταση:



Από τη γραφική παράσταση φαίνεται καθαρά ότι η ποσότητα e_3 είναι καλύτερη προσέγγιση του $\|e\|$ από το $\|r\|$.

Συνοπτικά τα χαρακτηριστικά των τεσσάρων πινάκων:

A	Διάσταση	$\kappa(A)$	$\ A\ $	$\frac{1}{\ A^{-1}\ }$	$\frac{1}{\kappa(A)}$
1	300×300	8.4×10^{11}	50	6×10^{-11}	1.2×10^{-12}
2	150×150	5.7×10^7	69	1.4×10^{-6}	2×10^{-8}
3	100×100	2.5×10^4	112	3.8×10^{-5}	4.3×10^{-3}
4	50×50	20	36	1.8	5×10^{-2}

Στη συνέχεια παρουσιάζω συγκεντρωτικά τα σχετικά σφάλματα της διαδικασίας 1. Ως σχετικό σφάλμα (ας είναι για τη ποσότητα $\|r\|$) επιλέγω τον τύπο (50 ο συνολικός αριθμός των δοκιμών):

$$\text{RelErr} = \frac{\sqrt{\sum_{i=1}^{50} (\|e\|_i - \|r\|_i)^2}}{\sum_{i=1}^{50} \|e\|_i} \cdot 100$$

Σχετικό σφάλμα	A_1	A_2	A_3	A_4
$\ r\ $	106	73	151	15.0
e_3	181	96	57	0.4

Συμπέρασμα των δοκιμών με τη διαδικασία 1: όταν ο δείκτης κατάστασης του πίνακα A είναι πολύ μεγάλος, το $\|r\|$ προσεγγίζει καλύτερα το σφάλμα $\|e\|$. Όσο ο δείκτης κατάστασης μειώνεται, το e_3 προσεγγίζει καλύτερα το σφάλμα.

Επειδή οι “μετρήσεις υπαίθρου” (οι μετρήσεις βαρύτητας και γενικά όλες οι γεωφυσικές μετρήσεις) είναι “μολυσμένες” με θόρυβο, προχωρώ στη **Διαδικασία 2**:

παίρνω όπως και στη διαδικασία 1 ένα τυχαίο μοντέλο και υπολογίζω το $g = Ax$ στο οποίο προσθέτω μια τυχαία μεταβλητή με μηδενικό μέσο όρο και μοναδιαία διακύμανση (η εντολή rand της Matlab έχει αυτή την ιδιότητα). Πολλαπλασιάζοντας κατάλληλα πετυχαίνω μια διαταραχή στις μετρήσεις μας με ένα επίπεδο θορύβου 2.5% (για τις μετρήσεις υπαίθρου θεωρείται ιδανική περίπτωση). Επίσης ξεκινάω από ένα τυχαίο μοντέλο $x^{(0)}$. Στη συνέχεια, όπως και στη διαδικασία 1 εκτελώ μια επανάληψη με το πρόγραμμα MLM.

Αλγόριθμος CheckErr2

Begin

for $i = 1, \dots, 50$

$$x = l + \text{rand} \cdot (u - l)$$

$$x^{(0)} = l + \text{rand} \cdot (u - l)$$

$$g = Ax$$

for $j = 1, \dots, n$

$$nse = 0.05 \cdot \text{rand} \cdot g_j - 0.025 \cdot g_j$$

$$g_j = g_j + nse$$

end for

$$x^{(1)} = \text{MLM}(x^{(0)})$$

$$g_c = Ax^{(1)}$$

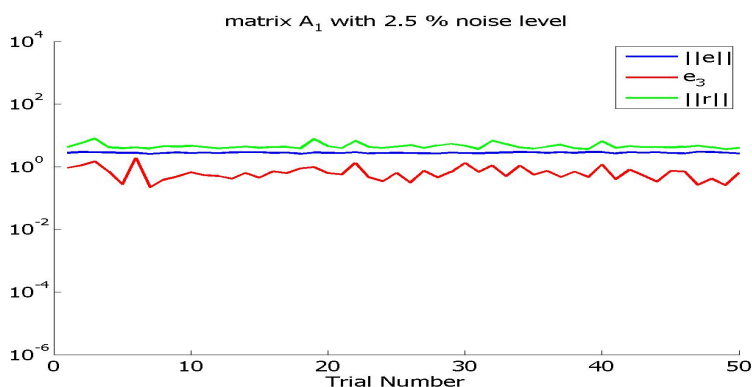
$$nE_i = \|x - x^{(1)}\|$$

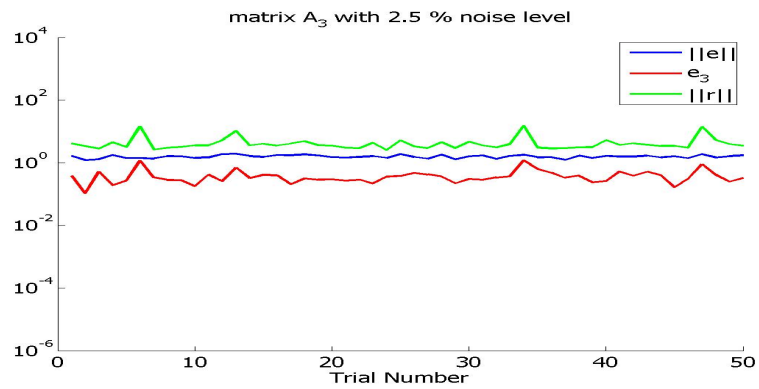
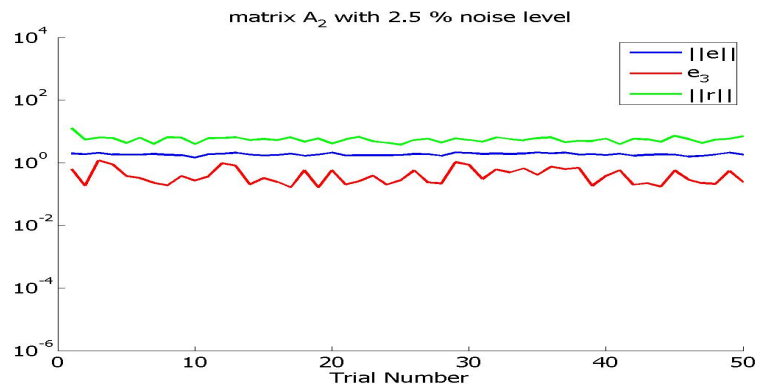
$$nE3_i = \text{fCalcErr}(g, g_c, A)$$

$$nR_i = \|g - g_c\|$$

end for

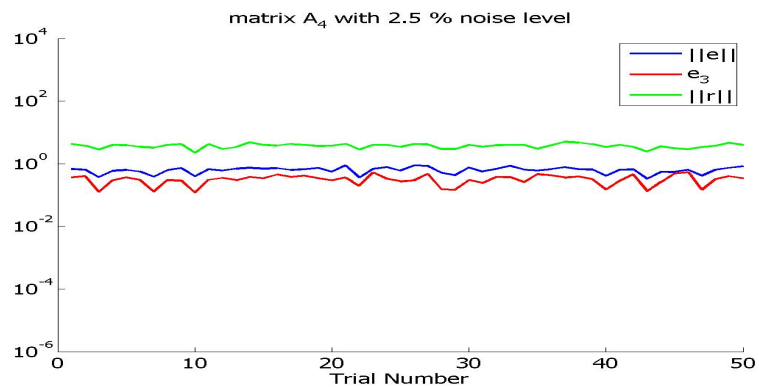
Παρουσιάζω τις γραφικές παραστάσεις για τους ίδιους 4 πίνακες:





και το συγκεντρωτικό πίνακα με τα σχετικά σφάλματα της διαδικασίας 2:

Σχετικό σφάλμα	A_1	A_2	A_3	A_4
$\ r\ $	142	550	1014	1078
e_3	165	111	98	28



Για επίπεδο θορύβου 5% παρουσιάζω μόνο το συγκεντρωτικό πίνακα:

Σχετικό σφάλμα	A_1	A_2	A_3	A_4
$\ r\ $	731	1896	2531	3045
e_3	190	135	105	42

Επειδή τα αρχικά μοντέλα είναι τυχαία, κάθε φορά που εκτελούμε τα προγράμματα τα αποτελέσματα των σχετικών σφαλμάτων είναι διαφορετικά. Έκανα πολλές δοκιμές και στους πίνακες έχω αναγράψει τους μέσους όρους των δοκιμών.

Γενικό συμπέρασμα: Όταν οι μετρήσεις μας περιέχουν θόρυβο, ανεξάρτητα από τη κατάσταση του πίνακα A , η εκτίμηση e_3 προσεγγίζει πολύ καλύτερα το σφάλμα $\|e\|$ από το $\|r\|$.

5.2 Συνάρτηση LSQLIN της Matlab

Σε όλες τις δοκιμές θα χρησιμοποιήσω τον πίνακα A_1 της προηγούμενης ενότητας με τον μεγαλύτερο δείκτη κατάστασης.

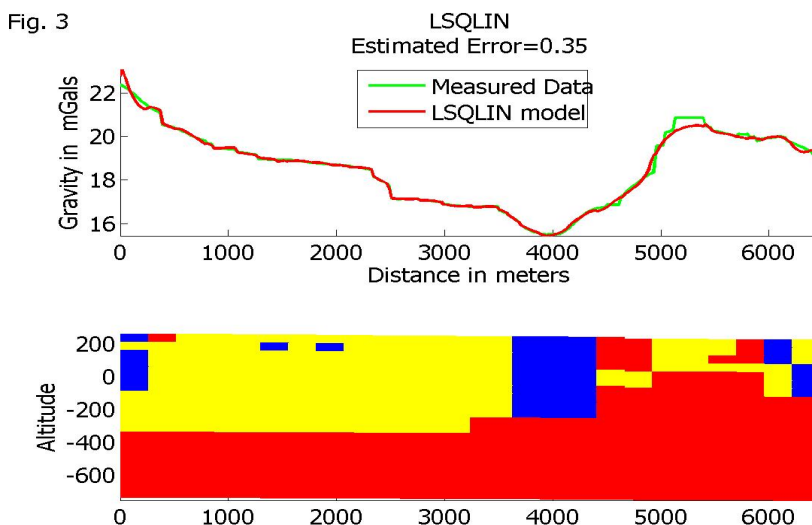
Η συνάρτηση LSQLIN της Matlab λύνει τα προβλήματα ελαχίστων τετραγώνων με υπό συνθήκη (constrained) λύση. Για το πρόβλημά μας, η εντολή γράφεται ως:

$$x = \text{lsqlin}(A, g, C, d, Ceq, deq, l, u, x0)$$

όπου οι πίνακες C, d, Ceq, deq χρησιμοποιούνται για άλλες ανισότητες και για την περίπτωση μας δηλώνονται ως κενοί.

Επιλέγουμε ως αρχικό μοντέλο $x_i^{(0)} = \frac{l_i + u_i}{2}, i = 1, \dots, n$.

Μετά από 36 επαναλήψεις, παίρνουμε τη γραφική παράσταση:



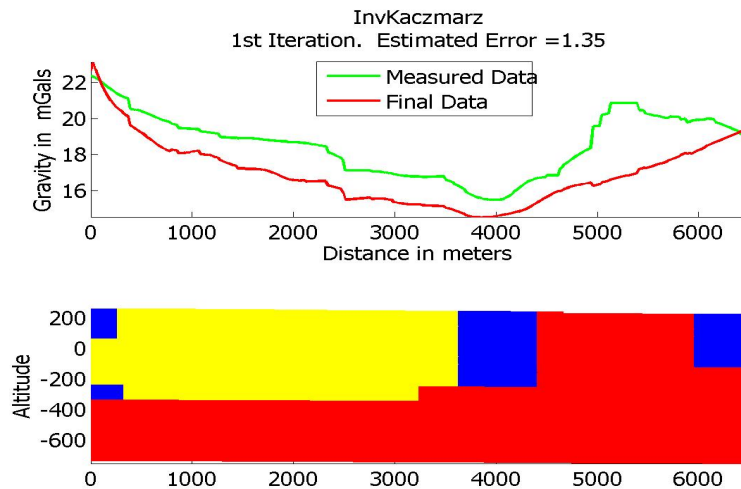
Στο παραπάνω σχήμα, με χρώμα **μπλέ** συμβολίζονται τα ιζύματα, χαλαρά επιφανειακά πετρώματα με πυκνότητες **2.00 - 2.25**. Με **κόκκινο** συμβολίζονται τα σκληρά πετρώματα (ασβεστόλιθος, γρανίτης) που αντιπροσωπεύουν το υπόβαθρο και έχουν πυκνότητες **2.50 - 2.70**. Με **κίτρινο** συμβολίζονται πετρώματα με ενδιάμεση πυκνότητα

2.25 - 2.50.

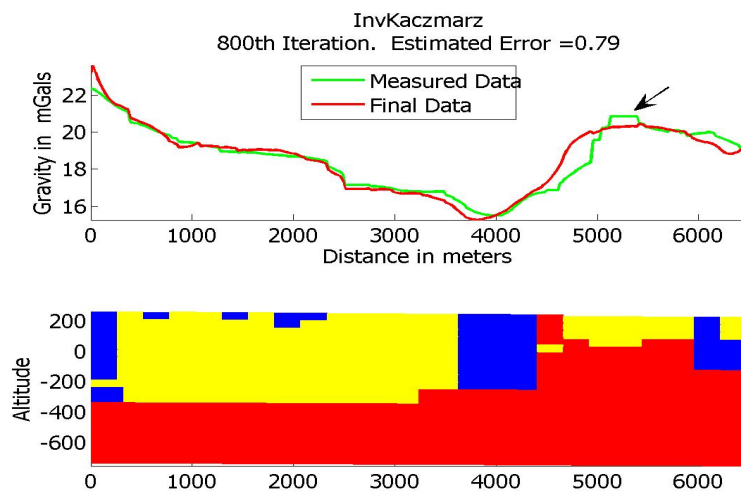
Παρατηρούμε μια καλή σύγκλιση ανάμεσα στις δύο καμπύλες αλλά το σφάλμα είναι αρκετά μεγάλο. Ας δούμε και άλλες μεθόδους.

5.3 Αλγόριθμος InvKaczmarz

Όπως στη προηγούμενη ενότητα, επιλέγουμε ως αρχικό μοντέλο $x_i^{(0)} = \frac{l_i + u_i}{2}$, $i=1, \dots, n$ και εκτελώ μια επανάληψη με το πρόγραμμα InvKaczmarz:



Μετά από 800 επαναλήψεις φτάνουμε στο αποτέλεσμα:

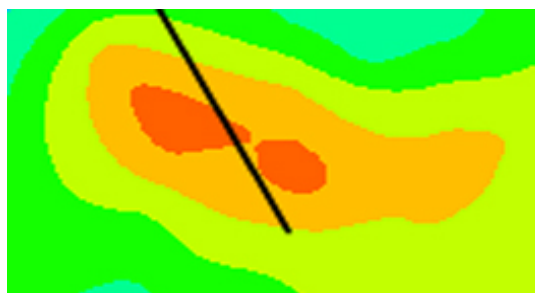


Παρατηρώ ότι υπάρχει μια παλινδρόμηση γύρω από τη βέλτιστη λύση. Κατά τη διάρκεια των επαναλήψεων, το

σφάλμα ανεβαίνει μέχρι την τιμή 11 για να ξανακατέβει, κτλ.

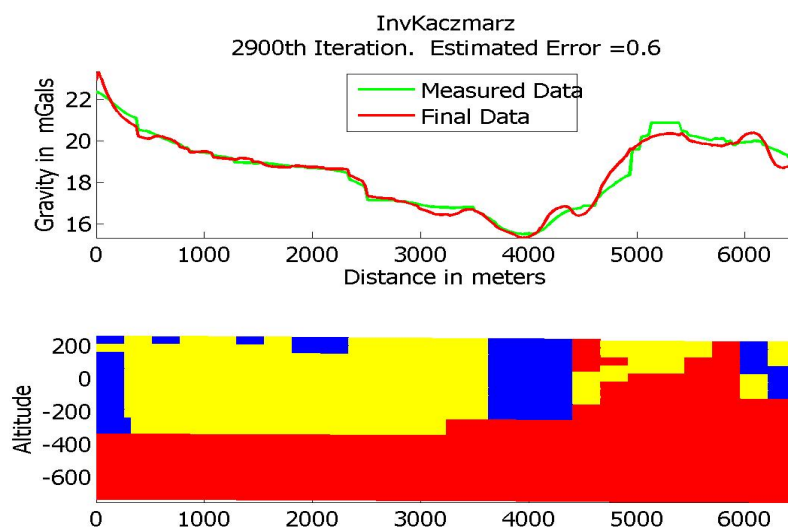
Τη γεωλογική ερμηνεία θα τη δούμε στο τέλος των δοκιμών. Για τη σύγκλιση, ενώ αριστερά έχουμε πολύ καλή, δεξιά μας τα χαλάει λίγο. Ο αλγόριθμος προσπαθεί να ερμηνεύσει την απότομη καμπυλότητα (πράσινη καμπύλη, στο βελάκι του άνω σχήματος) η οποία οφείλεται σε τρισδιάστατη ανωμαλία με δυσδιάστατο μοντέλο (ενότητα 2.3) και δεν έχουμε ικανοποιητικά αποτελέσματα.

Να δούμε τη σημαίνει τρισδιάστατη ανωμαλία.



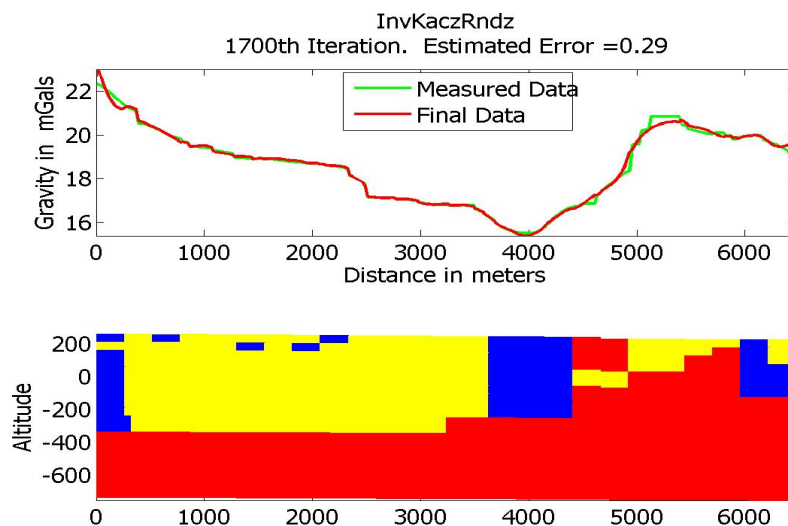
Παρατηρούμε στην πάνω εικόνα (η οποία είναι στο τέλος της A γραμμής) ότι η γραμμή μας επηρεάζεται από πλάγιες υψηλές τιμές (οι οποίες δημιουργούν ανωμαλίες) χωρίς να είναι ακριβώς πάνω στη γραμμή. Αυτό δημιουργεί μια "ψεύτικη" ανωμαλία η οποία δεν θα έπρεπε να ληφθεί υπόψη.

Μετά από 2900 επαναλήψεις παίρνουμε την ελάχιστη τιμή σφάλματος με το μοντέλο:



Με τον αλγόριθμο `InvKaczmarzRndz` μετά από 90 επαναλήψεις παίρνουμε σφάλμα $\text{estimated error } e_3 = 0.39$, αρκετά μικρότερο από του προηγούμενου αλγόριθμου.

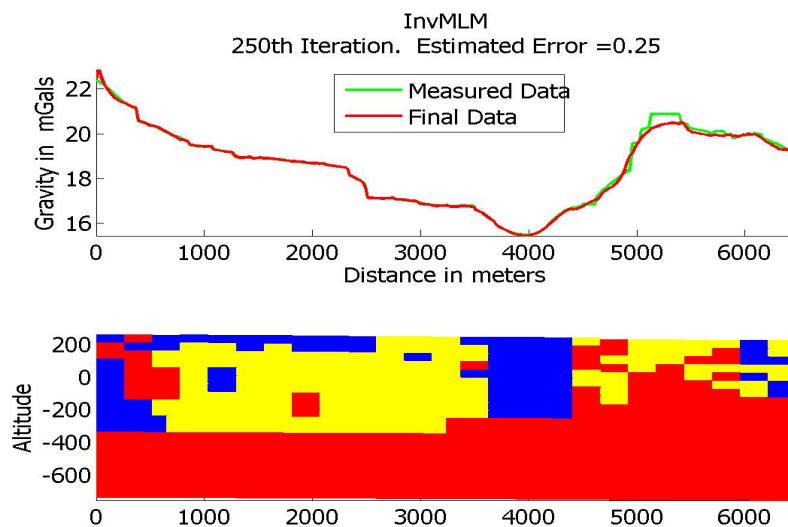
Μετά από 1700 επαναλήψεις:



Παρατηρούμε ότι έχουμε μια πολύ καλή σύγκλιση (σφάλμα 0.29, καλύτερο από το 0.35 της συνάρτησης L-SQLIN) άρα ο αλγόριθμος InvKaczmarzRndz δουλεύει αρκετά καλά. Επίσης, ο αλγόριθμος δεν προσπαθεί να ερμηνεύσει την απότομη κλίση η οποία οφείλεται σε τρισδιάστατη ανωμαλία, άρα έχει καλή συμπεριφορά.

5.4 Αλγόριθμος ModifiedLM

Επιλέγουμε το αρχικό μοντέλο όπως στη προηγούμενη ενότητα και εκτελούμε το πρόγραμμα invMLM. Μετά από 250 επαναλήψεις παίρνουμε το αποτέλεσμα:



Έχουμε αρκετά καλή σύγκλιση (σφάλμα 0.25 αντί 0.6 της μεθόδου InvKaczmarz και 0.29 της μεθόδου

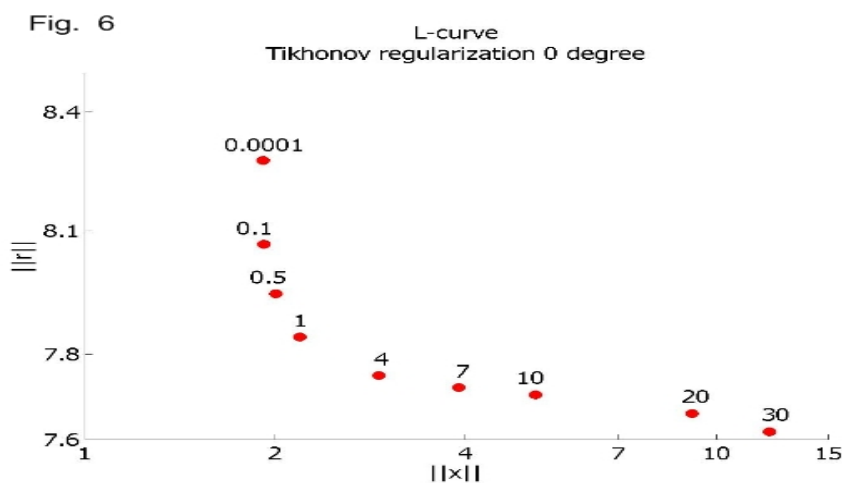
InvKaczmarzRndz) , εκτός από τη καμπυλότητα η οποία οφείλεται σε τρισδιάστατη ανωμαλία. Τελικά, δεν παίζει τόσο ρόλο το σφάλμα, περισσότερο μας ενδιαφέρει η σύγκλιση των δύο καμπυλών (μετρημένων τιμών και ανταπόκριση του μοντέλου) στη γραφική παράσταση. Για τα γεωφυσικά μοντέλα, ο αλγόριθμος MLM έχει καλύτερη συμπεριφορά, δεν προσπαθεί να ερμηνεύσει την καμπυλότητα που προέρχεται από τρισδιάστατη ανωμαλία, αλλά προσπαθεί να ακολουθήσει την κλήση της καμπύλης σε κάθε σημείο μέσα στα επιτρεπτά όρια (όπως είδαμε στην ενότητα 2.5, δεν επιτρέπονται απότομες κλήσεις στην καμπύλη των μετρήσεων υπαίθρου). Επίτηδες επέλεξα αυτή τη γραμμή για να δω τη συμπεριφορά των διαφόρων αλγορίθμων, αντί να επιλέξω μια λιγότερο απότομη γραμμή ή να την “ομαλοποιήσω” αυτήν.

5.5 Κανονικοποίηση Tichonov

Για την εκτίμηση της παραμέτρου λ για τη κανονικοποίηση Tikhonov μηδενικού βαθμού (εξίσωση 4.24) χρησιμοποίησα τη συνάρτηση Tikhstr με τις ακόλουθες δοκιμαστικές τιμές,

$$\lambda = [10^{-4}, 0.1, 0.5, 1, 4, 7, 10, 20, 30].$$

Για κάθε λ παίρνουμε μια νόρμα της λύσης $\|x\|$ και μια νόρμα των υπολοίπων $\|r\|$. Η γραφική παράσταση (σε λογαριθμική κλίμακα):

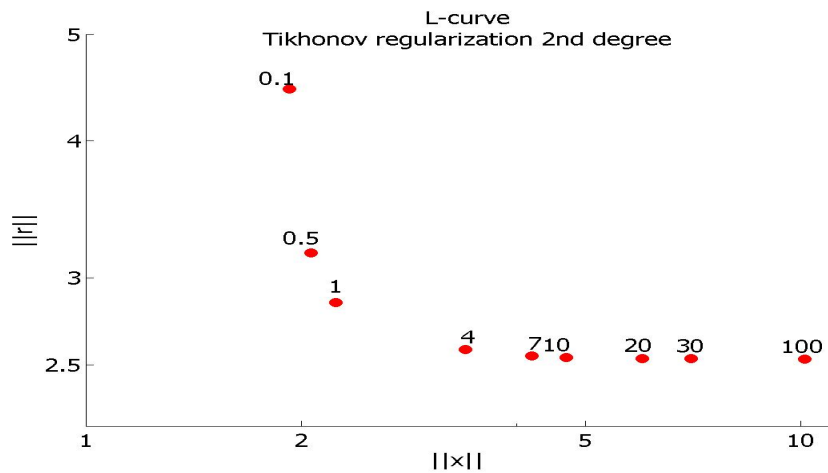
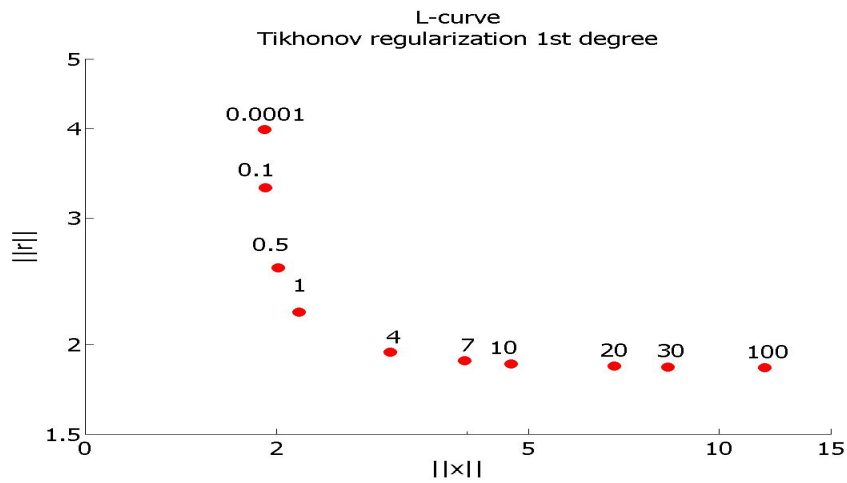


Παρατηρούμε τον σχηματισμό των δύο πλευρών, της κάθετης πλευράς στην οποία κυριαρχούν τα σφάλματα μέτρησης και της επίπεδης πλευράς στην οποία κυριαρχούν τα σφάλματα κανονικοποίησης, αλλά δεν έχουμε μια ευδιάκριτη γωνία η οποία να μας δίνει την τιμή της παραμέτρου λ .

Στην παρακάτω εικόνα, για τη κανονικοποίηση Tichonov πρώτου βαθμού, παρατηρούμε ότι η κάθετη πλευρά οριζοντιώνεται αλλά και πάλι δεν είναι ευδιάκριτη η γωνία η οποία ξεχωρίζει τις δύο μορφές σφαλμάτων.

Ας δούμε και την L-curve για την κανονικοποίηση Tichonov δεύτερου βαθμού:

Σε σχέση με τη κανονικοποίηση Tichonov πρώτου βαθμού δεν βελτιώνεται το αποτέλεσμα, άρα πρέπει να



βρούμε διαφορετικό τρόπο για την προσέγγιση της παραμέτρου λ .

Από την εξίσωση 4.20 παίρνουμε

$$\frac{\lambda^2 E x_\lambda}{A^T r_\lambda} = 1 \quad (5.2)$$

όπου $E = L^T L$ άρα για κάθε βαθμό κανονικοποίησης πρέπει να βρούμε την τιμή της παραμέτρου λ για την οποία ισχύει η 5.2 .

Τροποποίησα την συνάρτηση Golden της ενότητας 4.7 σε **Dw_Goldsec**, για να ελαχιστοποιήσω την ποσότητα $1 - \frac{\lambda^2 \|E x_\lambda\|}{\|A^T r_\lambda\|}$. Ο αλγόριθμος της αντίστοιχης συνάρτησης func θα είναι:

function $f = \text{func}(\lambda)$

% Μεταφορά στην μνήμη του πίνακα L του αντίστοιχου βαθμού κανονικοποίησης (ενότητα 4.4)

% Εκτελούμε το υποπρόγραμμα Tickerstr με τους περιορισμούς και παίρνουμε το διάνυσμα x

% Υπολογίζουμε τις ποσότητες $g_c, r_\lambda = g - g_c$ και στη συνέχεια την $f = 1 - \frac{\lambda^2 \|Ex_\lambda\|}{\|A^T r_\lambda\|}$.

Κανονικοποίηση Tichonov βαθμού	λ
0	1.171
1	2.068
2	1.052

Θα παρουσιάσω τα αποτελέσματα άλλων δύο μεθόδων εκτίμησης της παραμέτρου λ .

Ο P. C. Hansen στην εργασία του "The L-curve and its use in the numerical treatment of inverse problems" αναφέρει το **Κριτήριο Μέγιστης Καμπυλότητας** ανάμεσα στις ποσότητες $\eta = \|x_\lambda\|^2$, $\rho = \|Ax_\lambda - b\|^2$ όπου x_λ είναι μια λύση του συστήματος 4.18 για κάποια τιμή της παραμέτρου λ . Μετά από εφαρμογή της μεθόδου (συνδυάζοντας και την συνάρτηση Golden της ενότητας 4.7) είχαμε τα αποτελέσματα

Κανονικοποίηση Tichonov βαθμού	λ
0	1.107
1	1.296
2	1.075

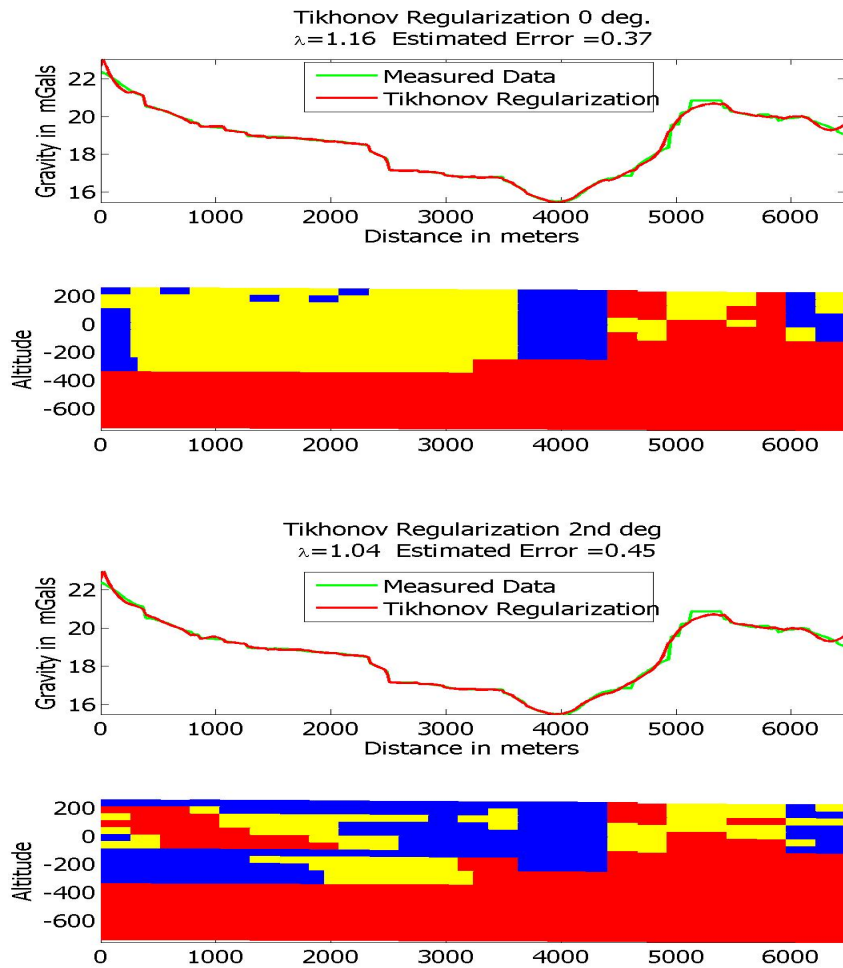
Η μέθοδος "Generalized cross validation" [18] ακολουθεί την διαδικασία: από το σύνολο των n εξισώσεων της μορφής $A_{i,:}x_i = g_i$, $i = 1, \dots, n$ αφίρουμε k απ' έξω και υπολογίζουμε τις παραμέτρους x_i , $i = 1, \dots, n-k$ για διάφορες τιμές της παραμέτρου λ . Στη συνέχεια, για κάθε τιμή της παραμέτρου λ με γραμμική παρεμβολή υπολογίζουμε τις τιμές των υπολοίπων k παραμέτρων x_{n-k+1}, \dots, x_n και παίρνουμε τη νόρμα των διαφορών από τα πραγματικά. Μας ενδιαφέρει η τιμή της παραμέτρου λ για το οποίο η νόρμα είναι ελάχιστη.

Κανονικοποίηση Tichonov βαθμού	λ
0	$< 10^{-4}$
1	0.252
2	0.150

Η βιβλιογραφία που βρήκα για την τελευταία μέθοδο αφορούν λύσεις του συστήματος 4.18 χωρίς περιορισμούς. Στο διαδίκτυο δεν βρήκα κανένα άρθρο για περιορισμούς στις λύσεις, για αυτό συνδύασα σε ένα πρόγραμμα την λογική της μεθόδου με τη συνάρτηση Tickerstr που συναντήσαμε στην ενότητα 4.5.

Συμπέρασμα, από τις δύο πρώτες μεθόδους, παίρνουμε παραπλήσιες τιμές για την παράμετρο λ για τη κανονικοποίηση μηδενικού και δεύτερου βαθμού, άρα βλέπουμε τα αντίστοιχα μοντέλα

Η κανονικοποίηση Tichonov δεύτερου βαθμού έχει μεγαλύτερο σφάλμα από την αντίστοιχη του μηδενικού βαθμού. Η κανονικοποίηση δεύτερου βαθμού όμως έχει καλύτερη συμπεριφορά, μοιάζει με τη συμπεριφορά της μεθόδου MLM, δεν προσπαθεί να ταιριάζει τις καμπύλες στο σημείο της απότομης κλήσης που οφείλεται

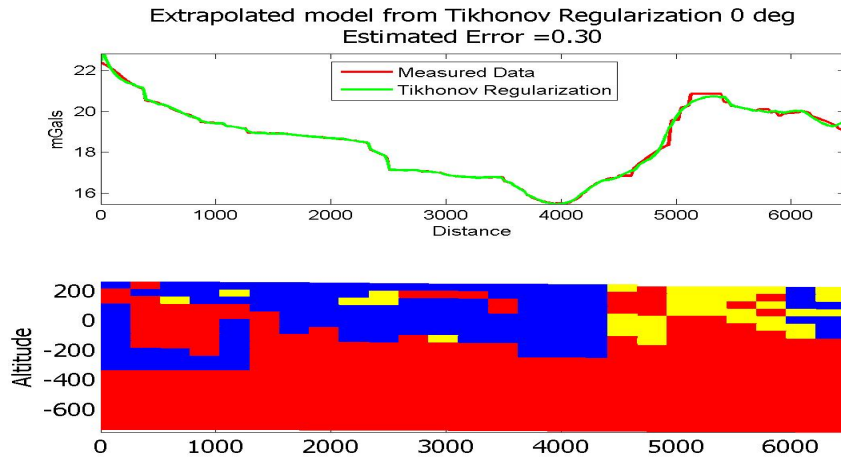


σε τρισδιάστατη ανωμαλία, ενώ η κανονικοποίηση μηδενικού βαθμού ταιριάζει με τη συμπεριφορά της μεθόδου InvKaczmarz. Επίσης, από μια μέθοδο κανονικοποίησης δεν ζητάμε να ακολουθεί κάθε “σπάσιμο” της καμπύλης, ζητάμε μια εξομάλυνση (smooth) των δεδομένων μας και αυτό το κάνει καλύτερα η κανονικοποίηση Tikhonov δεύτερου βαθμού. Αν ζητάμε ακριβή σύγκλιση στα δεδομένα ή εάν τα αρχικά δεδομένα έχουν εξομαλυνθεί με κάποια φίλτρα, θα ταιριάζει περισσότερο η κανονικοποίηση Tikhonov μηδενικού βαθμού.

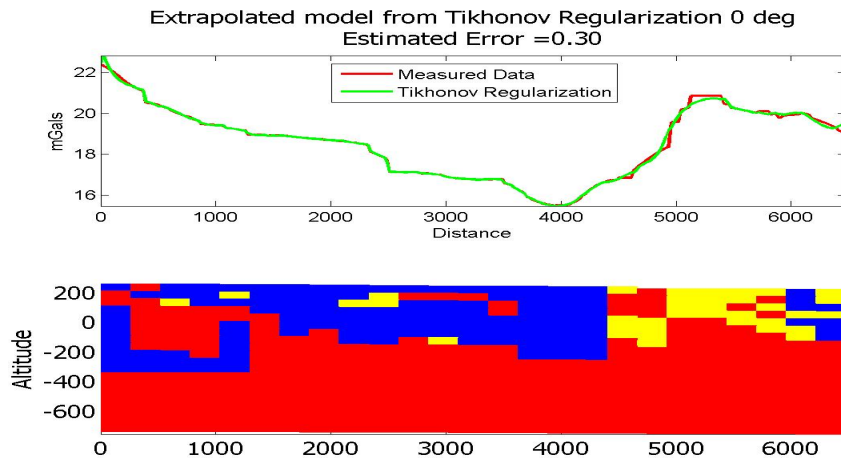
5.6 Τεχνικές παρεκβολής για γραμμικά συστήματα σε κακή κατάσταση

Για τη κανονικοποίηση Tikhonov μηδενικού βαθμού κάλεσα τη συνάρτηση Tikhonstr με $\lambda_1 = 10^{-5}$ και πήρα ως αποτέλεσμα το διάνυσμα \bar{x}_1 . Επίσης με $\lambda_2 = 1.14$ πήρα το \bar{x}_2 . Επέλεξα $\lambda_1 = 10^{-5}$ για να είναι κοντά στο μηδέν (όχι πιο μικρό ώστε να έχουμε σφάλματα κανονικοποίησης) και $\lambda_2 = 1.14$, η προσέγγιση που πήραμε από της μεθόδους της προηγούμενης ενότητας.

Τώρα μπορούμε να υπολογίσουμε τους συντελεστές $a_j, b_j, j = 1, \dots, n$ των εξισώσεων 4.32, 4.33 και στη συνέχεια τη κανονικοποιημένη λύση 4.28 .



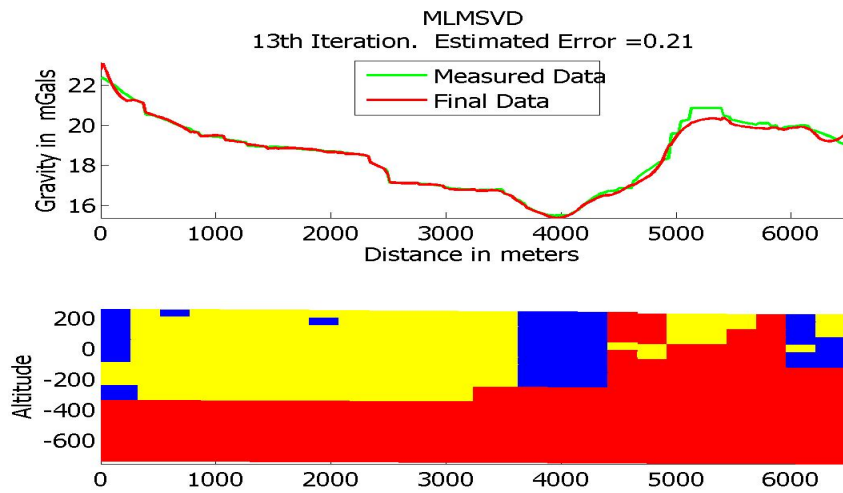
Ακολουθήσα την ίδια τακτική για τη κανονικοποίηση Tikhonov δεύτερου βαθμού για να έχουμε την λύση:



Παρατηρούμε ίδια αποτελέσματα (η νόρμα των διαφορών των δύο μεθόδων είναι της τάξης του 10^{-4}), πολύ κοντινά σφάλματα άρα η μέθοδος ανταποκρίνεται. Επίσης, το γεωλογικό μοντέλο (η κατανομή και το μέγεθος των πυκνοτήτων) μοιάζει αρκετά με τη κανονικοποίηση Tikhonov δεύτερου βαθμού.

5.7 Λύση MLM με SVD

Μετά από 13^η επαναλήψεις παίρνουμε τα αποτελέσματα:

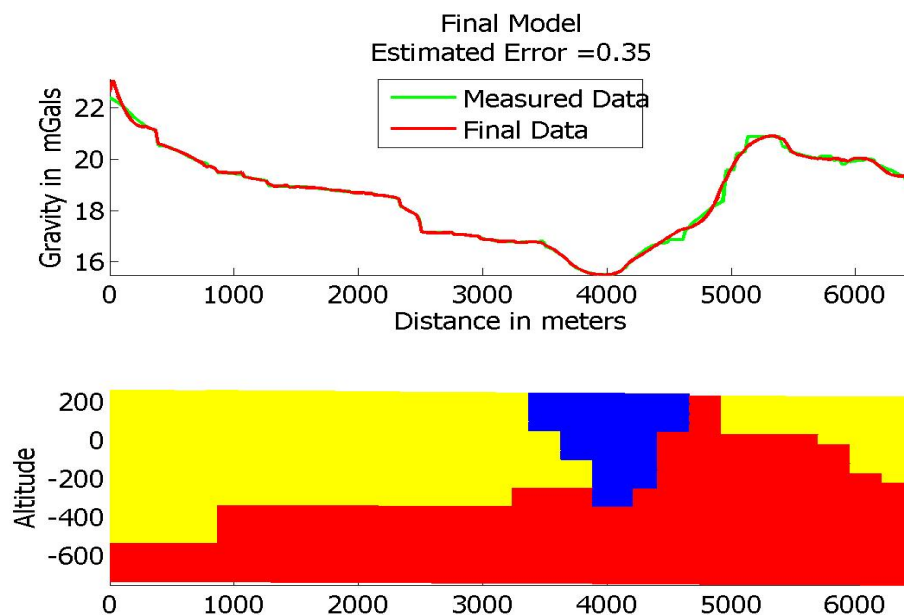


Ας δούμε ένα συγκεντρωτικό πίνακα με όλες τις μεθόδους:

Μέθοδος	Estim. Err. e_3	No. of iteration	flops/iter
lsqlin (MATLAB)	0.35	36	
InvKaczmarz	0.60	2900	$3n$
InvKaczmarzRnd	0.29	1700	$3n$
invMLM	0.25	250	$n^3/3$
Tikh. 0 deg	0.37		
Tikh. 2 deg	0.45		
Extrapol	0.30		
MLMSVD	0.21	13	$2n^2$

6 Γεωλογική Ερμηνεία

Μελετώντας τα αποτελέσματα όλων των δοκιμών, καταλήξαμε στο τελικό μοντέλο, το οποίο είναι κοντά στη γεωλογική ερμηνεία:



Βλέπουμε καθαρά το ρήγμα (μπλε χρώμα) το οποίο έχει δημιουργηθεί στην επαφή δύο διαφορετικών στρωμάτων: των χαλαρών ιζημάτων (κίτρινο χρώμα) και του ασβεστόλιθου (κόκκινο χρώμα). Το ρήγμα το οποίο η γραμμή μας το κόβει, είναι αποτυπωμένο σε γεωλογικούς χάρτες από γεωλογικές μελέτες. Η γεωλογική μελέτη πραγματοποιείται από παρατηρήσεις των επιφανειακών πετρωμάτων, μελέτη των κλήσεών τους, διαφοροποίησή τους από σημείο σε σημείο, κτλ. Θέλαμε να διακρίνουμε το ρήγμα και με γεωφυσικές μετρήσεις, επιβεβαιώθηκε, άρα η δουλειά μας ήταν σωστή. Επίσης, με τη γεωφυσική μελέτη εκτιμήσαμε το βάθος του ρήγματος, περίπου 550 μέτρα, εκτίμηση την οποία η γεωλογική μελέτη δεν μπορεί να την κάνει. Γι' αυτό χρειάζονται και οι δύο μελέτες (η γεωλογική οδήγησε στην γεωφυσική), η γεωλογική μελέτη γίνεται σε συντομότερο χρόνο και καλύπτει μεγάλο χώρο σε αντίθεση με την γεωφυσική μελέτη η οποία είναι περισσότερο τοπική και χρονοβόρα.

Αναφορές

- [1] B. Banerjee, 1977, "Gravitational attraction of a rectangular parallelepiped", *Geophysics* 42, 1053 (1977); doi:10.1190/1.1440766
- [2] C. Brezinski - G. Rodriguez - S. Seatsu, 2008, "Error estimates for linear systems with application to regularization, Feb 2008", *Numer. Algor.* (2008) 49:85-104, DOI 10.1007/s11075-008-9163-1
- [3] C. Brezinski - G. Rodriguez - S. Seatsu, 2008, "Error estimates for the regularization of least squares problem, Oct 2008", *Numer. Algor.* (2009) 51:61-76, DOI 10.1007/s11075-008-9243-2
- [4] C. Brezinski, 1999, "Error estimates for the solution of linear systems", *SIAM J. SCI. COMPUT.*, Vol. 21, No. 2, pp. 764-781
- [5] K. Levenberg, 1944, "A Method for the Solution of Certain Problems in Least-Squares", *Quart. Appl. Math.* 2, pp. 164-168, 1944
- [6] D. Marquardt, 1963, "An Algorithm for Least-Squares Estimation of Nonlinear Parameters", *SIAM J. Appl. Math.* 11, pp. 431-441, 1963
- [7] G. Auchmuty, 1992, "A posterior error estimates for linear equations.", *Numer. Math.* 61, 1-6, 1992
- [8] S. Kaczmarz, 1937, "Angenäherte Auflösung von Systemen linear Gleichungen", *Bull. Internat. Acad. Polon. Sci. Lettres A*, pages 335-357
- [9] W.H. Press, B. P. Flannery, S.A. Teukolsky, W.T. Vetterling, 1986, "Numerical Recipes, the Art of Scientific Computing", Book
- [10] T. Strohmer, R. Vershynin, "A randomized Kaczmarz algorithm with exponential convergence", *Dep. of Math., Univ. of California Davis, CA 95616-8633, USA*
- [11] P. C. Hansen, D. P. O'Leary, 1993, "The use of L-Curve in the regularization of discrete ill-posed problems", *SIAM J. Sci. Comput.*, Vol. 14, No. 6, pp1487-1503, Nov. 1993
- [12] A. N. Tikhonov, V. Y. Arsenin, 1977, "Solutions of Ill-Posed Problems", Wiley, New York, 1977
- [13] J. Skilling, S. Gull, 1985, "Algorithms and applications in Maximum-Entropy and Bayesian Methods in Inverse Problems", C. R. Smith and W. T. Grandy. Jr., eds., D. Reidel Pub. Co., Boston 1985, pp. 83-132
- [14] P. C. Hansen, 1992, "Analysis of discrete ill-posed problems by means of the L-Curve", *SIAM Rev.*, 34 (1992), pp. 561-580
- [15] C. L. Lawson, R.J. Hanson, 1974, "Solving Least Square Problems", Prentice-Hall, Englewood Cliffs, NJ, 1974
- [16] P. C. Hansen, 1990, "Relations between SVD and GSVD of discrete regularization problems in standard and general form", *Linear Algebra and its Application*, Vol. 141, Nov. 1990, pp. 165-176; doi:10.1016/0024-3795(90)

- [17] R. C. Aster, B. Borchers, C. H. Thurber, 2005, "Parameter Estimation and the Inverse Problems", Elsevier Academic Press
- [18] G. H. Golub, M. Heath, G. Wahba, 1979, "Generalized cross-validation as a method for choosing a good ridge parameter", *Technometrics* 21: 215-223
- [19] C. Brezinski, M. Redivo-Zaglia, G. Rodriguez, S. Seatzu, "Extrapolation techniques for ill-conditioned linear systems", *Numer. Math.* (1998) 81:1-29
- [20] M. Talwani, J. L. Worzel, M. Landisman (1959), "Rapid Gravity Computations for two-dimensional bodies with application to the Mendocine Submarine Fracture Zone", *J. Geophysical Research* 64: pp 49-61