



Διακριτά Αντίστροφα Προβλήματα 2021-2022

Διάλεξη 10

Παραμετροποίηση – Μη γραμμικά αντίστροφα
προβλήματα

Μιχάλης Ταρουδάκης

Παραμετροποίηση

$$f(\mathbf{m}, \mathbf{d}) = 0$$

Είναι σημαντικό στη διατύπωση ενός αντίστροφου προβλήματος να επιλέξουμε την κατάλληλη ομάδα παραμέτρων που θα έχουν βέβαια νόημα στην επίλυση ενός προβλήματος αλλά και την κατάλληλη ομάδα μετρήσεων

Μερικές φορές το ίδιο το πρόβλημα οδηγεί σε συγκεκριμένη ομάδα μετρήσεων ή παραμέτρων, σε άλλες όμως περιπτώσεις υπάρχει επιλογή.

Έχουμε δει τη δυνατότητα μετασχηματισμού παραμέτρων και μετρήσεων. Οι λύσεις παραμένουν ανεξάρτητες του μετασχηματισμού μόνο στην περίπτωση γραμμικών αντιστρόφων προβλημάτων με γκαουσιανή στατιστική

Παραμετροποίηση

$$\mathbf{f}(\mathbf{m}, \mathbf{d}) = 0$$

Σημασία παραμετροποίησης

Παράδειγμα ευθείας

$$d_i, i = 1, \dots, N \quad (z_i, d_i) \quad z_i, i = 1, \dots, N$$

Υπολογισμός 2 παραμέτρων από N μετρήσεις

(1,1), (2,2), (3,3), (4,5).

Με λύση ελαχίστων τετραγώνων

$$d = -0.5 + 1.3z$$

Παραμετροποίηση

$$f(\mathbf{m}, \mathbf{d}) = 0$$

Παράδειγμα ευθείας

(1,1), (2,2), (3,3), (4,5).

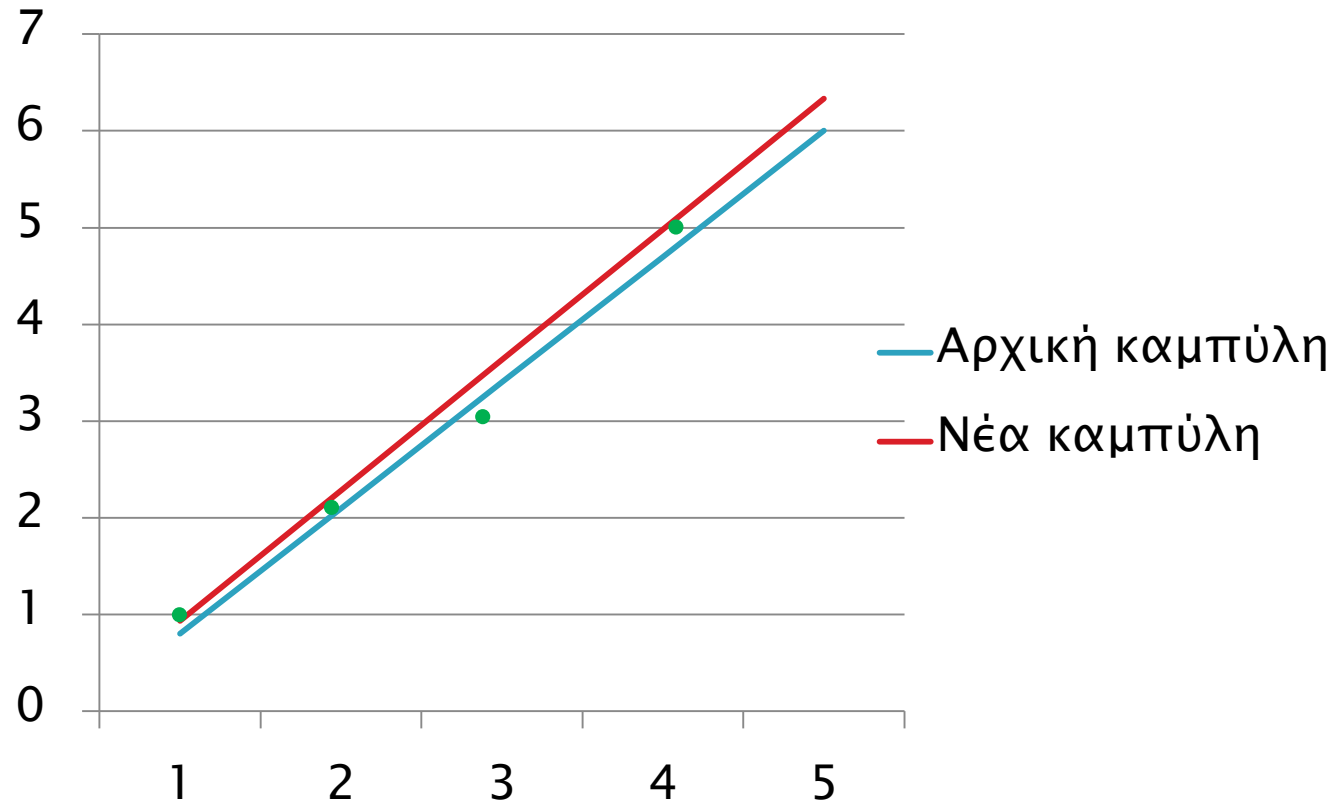
$$z \rightarrow d' \quad d \rightarrow z'$$

Με λύση ελαχίστων τετραγώνων $d' = -0.309 + 0.743z'$

$$z' = -0.416 + 1.34d' \quad d = -0.5 + 1.3z$$

Διαφορετικές εξισώσεις

Παραμετροποίηση



Παραμετροποίηση

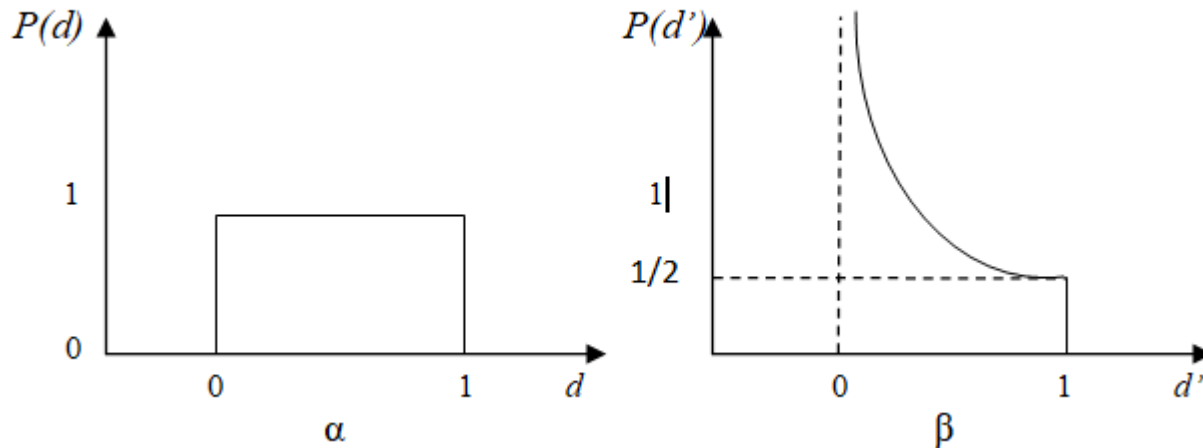
Που οφείλεται η διαφορά ;

Όταν θεωρούμε ως ακριβή την τετμημένη, η λύση ελαχίστων τετραγώνων θεωρεί ότι οι μετρήσεις ακολουθούν κανονική κατανομή.

Αλλάζοντας τετμημένες με τεταγμένες, θεωρούμε ότι οι δεύτερες είναι ακριβείς ενώ οι τετμημένες ακολουθούν κανονική κατανομή.

Εάν χρησιμοποιήσουμε κατανομές πιθανοτήτων, αλλαγή της παραμετροποίησης οδηγεί σε διαφορετικές εκτιμήσεις, παρά το γεγονός ότι κατανομές πιθανοτήτων είναι αναλλοίωτες σε αλλαγές παραμετροποίησης.

Παραμετροποίηση

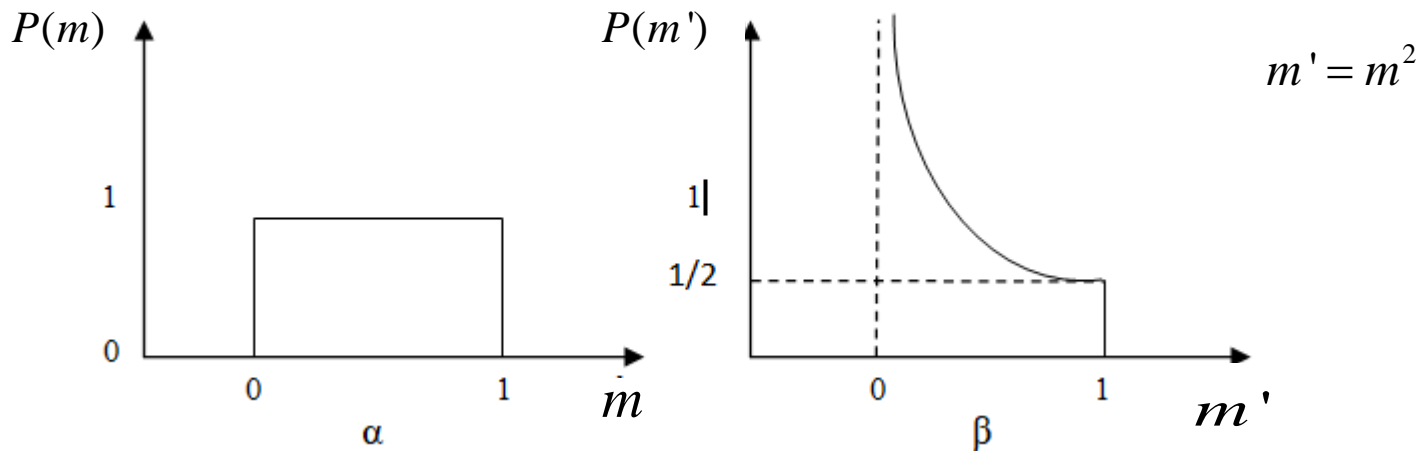


$$d' = d^2$$

$$P(d)\partial d = P(d')\partial d'$$

$$P(d') = P(d) \frac{\partial d}{\partial d'} = 1 \cdot \frac{1}{2\sqrt{d'}}$$

Παραμετροποίηση



Η m δεν έχει λύση μεγίστου. Η m' έχει λύση μεγίστου στο $m' = 0$

$$E[m] = \int_0^1 m P(m) \partial m = \int_0^1 m \partial m = \frac{1}{2} \quad E[m'] = \int_0^1 m' P(m') \partial m' = \int_0^1 \frac{1}{2} \sqrt{m'} \partial m' = \frac{1}{3}$$

$$\frac{1}{3} \neq \left(\frac{1}{2}\right)^2$$

Παραμετροποίηση

Ο λόγος που αλλάζουμε μερικές φορές παραμετροποίηση είναι γιατί μπορεί να μετατρέψουμε ένα αντίστροφο πρόβλημα σε κάποιο ισοδύναμο που μπορεί να λυθεί με κάποιο πιο εύκολο τρόπο

Παραμετροποίηση

$$d_i = m_1 \exp(m_2 z_i)$$

$$m_1' = \ln(m_1), \quad m_2' = m_2 \quad \text{και} \quad d_i' = \ln(d_i)$$

$$d_i' = m_1' + m_2' z_i$$

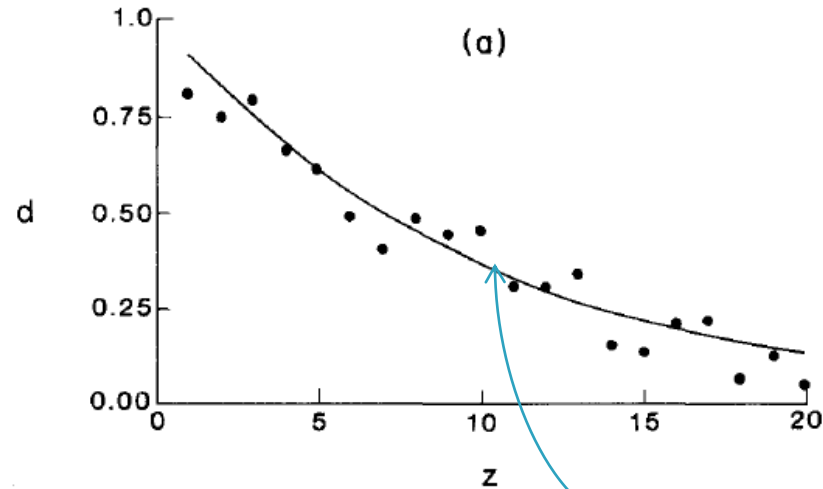
Λύση ελαχίστων τετραγώνων :

Οι μετρήσεις d_i' υπόκεινται σε λάθη που ακολουθούν κανονική κατανομή.

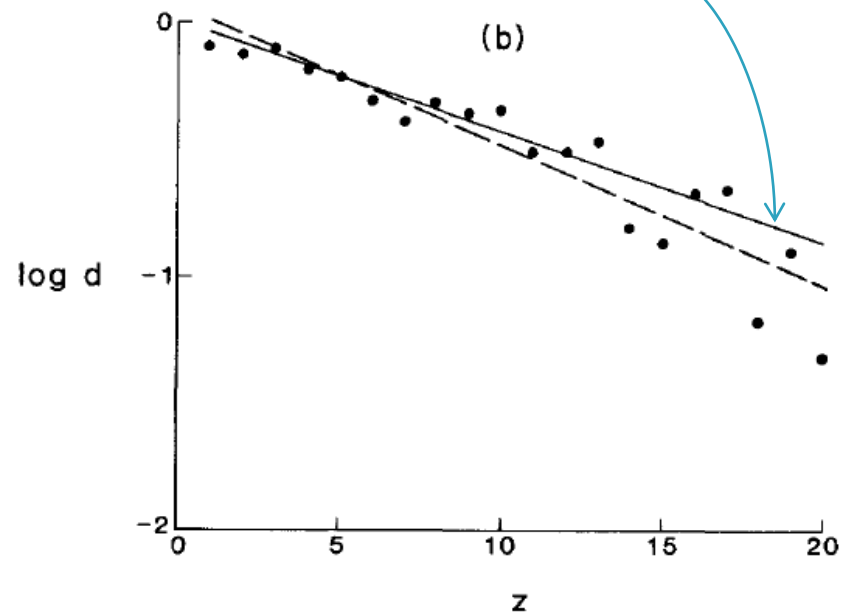
Προσοχή όμως. Η τελευταία υπόθεση συνεπάγεται ότι οι αρχικές μετρήσεις δεν πρέπει να ακολουθούν κανονική κατανομή !

Παραμετροποίηση

Οι μετρήσεις ακολουθούν κανονική κατανομή στην αρχική παραμετροποίηση



Η νέα παραμετροποίηση ενισχύει τη διασπορά των μετρήσεων σε μεγάλες τετμημένες



Μη γραμμικά αντίστροφα προβλήματα

Γενικό αντίστροφο πρόβλημα

$$\mathbf{f}(\mathbf{m}, \mathbf{d}) = 0$$

Γραμμικό αντίστροφο πρόβλημα

$$\mathbf{G}\mathbf{m} = \mathbf{d}$$

Θεωρούμε ότι τα δεδομένα μας ακολουθούν κανονική κατανομή με γνωστό μέσο και αντίστοιχα έχουμε αρχικές πληροφορίες για τις παραμέτρους που επίσης ακολουθούν κανονική κατανομή.

$$\mathbf{x} = [\mathbf{d}, \mathbf{m}]^T$$

Η αρχική πληροφορία για μετρήσεις και παραμέτρους βρίσκονται σε ένα «νέφος» (cloud) $S(\mathbf{x})$ γύρω από τις μετρήσεις και μέσες τιμές παραμετρών το σχήμα του οποίου καθορίζεται από τον πίνακα συνδιακύμανσης $[\text{cov}\mathbf{x}]$

Μη γραμμικά αντίστροφα προβλήματα

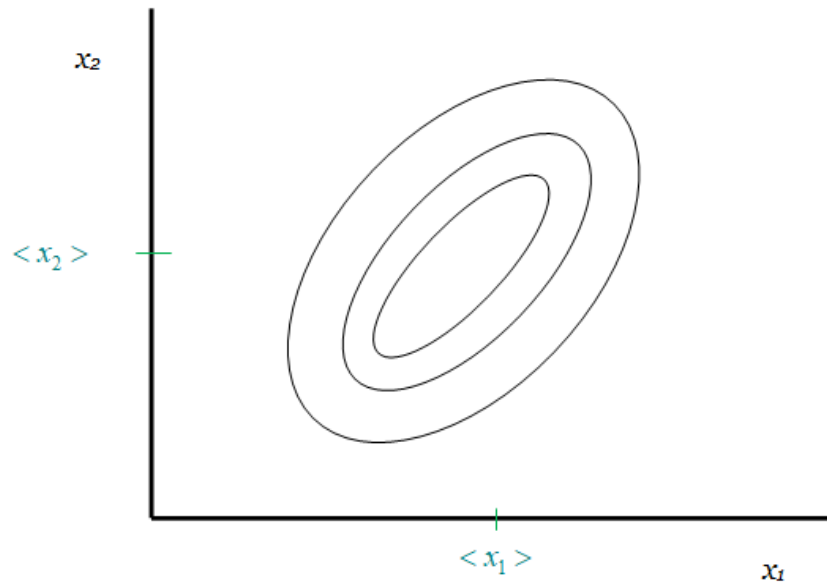
Ο πίνακας συνδιακύμανσης $[\text{cov} \mathbf{x}]$ περιλαμβάνει τους πίνακες συνδιακύμανσης μετρήσεων και παραμέτρων σε διαγώνια blocks

$$\mathbf{x} = [\mathbf{d}, \mathbf{m}]^T$$

$$P_A(\mathbf{x}) \propto \exp\left[-\frac{1}{2} (\mathbf{x} - \langle \mathbf{x} \rangle)^T [\text{cov} \mathbf{x}]^{-1} (\mathbf{x} - \langle \mathbf{x} \rangle)\right]$$

$$\langle \mathbf{x} \rangle = [\mathbf{d}^{obs}, \langle \mathbf{m} \rangle]^T$$

Μη γραμμικά αντίστροφα προβλήματα

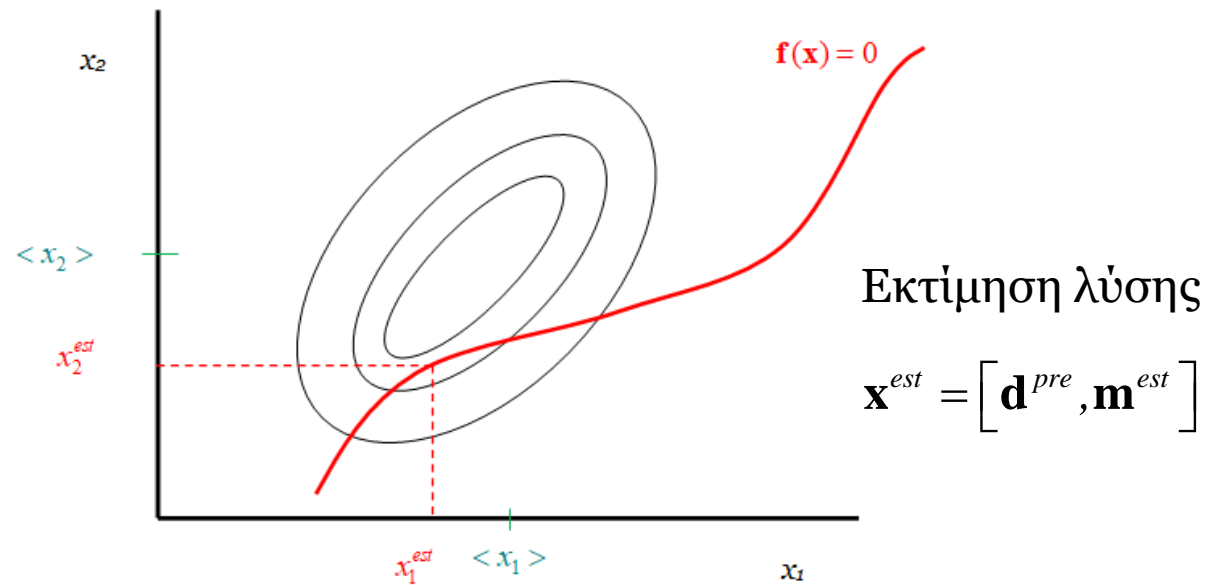


Ο χώρος $S(\mathbf{x})$ έχει
διάσταση $N+M$

$$\langle \mathbf{x} \rangle = [\mathbf{d}^{obs}, \langle \mathbf{m} \rangle]^T$$

Μη γραμμικά αντίστροφα προβλήματα

Ακριβές μοντέλο $\mathbf{f}(\mathbf{x})=0$



$$P_A(\mathbf{x}) \propto \exp\left[-\frac{1}{2}(\mathbf{x} - \langle \mathbf{x} \rangle)^T [\text{cov } \mathbf{x}]^{-1}(\mathbf{x} - \langle \mathbf{x} \rangle)\right]$$

Ελαχιστοποίηση

$$\Phi = [(\mathbf{x} - \langle \mathbf{x} \rangle)^T [\text{cov } \mathbf{x}]^{-1}(\mathbf{x} - \langle \mathbf{x} \rangle)]$$

Μη γραμμικά αντίστροφα προβλήματα

Πολλαπλασιαστές Lagrange

$$\partial\Phi / \partial x_i = \sum_{j=1}^p 2\lambda_j \partial f_j / \partial x_i$$

$$[\mathbf{x} - \langle \mathbf{x} \rangle]^T [\text{cov } \mathbf{x}]^{-1} = \boldsymbol{\lambda}^T \mathbf{F}$$

όπου $\boldsymbol{\lambda}$ είναι το διάνυσμα των πολλαπλασιαστών Lagrange και \mathbf{F} είναι ο πίνακας των κλίσεων $\nabla \mathbf{f}$

Μη γραμμικά αντίστροφα προβλήματα

Πολλαπλασιαστές Lagrange

$$[\mathbf{x} - \langle \mathbf{x} \rangle]^T [\text{cov } \mathbf{x}]^{-1} = \boldsymbol{\lambda}^T \mathbf{F}$$

$$[\mathbf{x} - \langle \mathbf{x} \rangle]^T = [\text{cov } \mathbf{x}] \boldsymbol{\lambda}^T \mathbf{F}$$

$$[\mathbf{x} - \langle \mathbf{x} \rangle] = [\text{cov } \mathbf{x}] \mathbf{F}^T \boldsymbol{\lambda}$$

$$\mathbf{F}[\mathbf{x} - \langle \mathbf{x} \rangle] = \mathbf{F}[\text{cov } \mathbf{x}] \mathbf{F}^T \boldsymbol{\lambda}$$

$$\boldsymbol{\lambda} = (\mathbf{F}[\text{cov } \mathbf{x}] \mathbf{F}^T)^{-1} \mathbf{F}[\mathbf{x} - \langle \mathbf{x} \rangle]$$

Μη γραμμικά αντίστροφα προβλήματα

Πολλαπλασιαστές Lagrange

$$[\mathbf{x} - \langle \mathbf{x} \rangle]^T [\text{cov } \mathbf{x}]^{-1} = \boldsymbol{\lambda}^T \mathbf{F}$$

$$\boldsymbol{\lambda} = (\mathbf{F}[\text{cov } \mathbf{x}]\mathbf{F}^T)^{-1} \mathbf{F}[\mathbf{x} - \langle \mathbf{x} \rangle]$$

$$[\mathbf{x} - \langle \mathbf{x} \rangle] = [\text{cov } \mathbf{x}]\mathbf{F}^T (\mathbf{F}[\text{cov } \mathbf{x}]\mathbf{F}^T)^{-1} \mathbf{F}[\mathbf{x} - \langle \mathbf{x} \rangle]$$

Επίλυση ταυτόχρονα με $\mathbf{f}(\mathbf{x})=0$

$$[\mathbf{x} - \langle \mathbf{x} \rangle] = [\text{cov } \mathbf{x}]\mathbf{F}^T (\mathbf{F}[\text{cov } \mathbf{x}]\mathbf{F}^T)^{-1} \mathbf{F}[\mathbf{x} - \langle \mathbf{x} \rangle] - \mathbf{f}(\mathbf{x})$$

Μη γραμμική εξίσωση – Δεν επιλύεται εύκολα

Μία επαναληπτική μέθοδος

Υπόθεση :

\mathbf{x}_0^{est} πολύ κοντά στη λύση μέγιστης πιθανοφάνειας

$$\mathbf{x}_{n+1}^{est} = \langle \mathbf{x} \rangle + [\text{cov } \mathbf{x}] \mathbf{F}_n^T (\mathbf{F}_n [\text{cov } \mathbf{x}] \mathbf{F}_n^T)^{-1} \mathbf{F}_n [\mathbf{x}_n^{est} - \langle \mathbf{x} \rangle] - \mathbf{f}(\mathbf{x}_n^{est})$$

Περίπτωση άμεσης μη γραμμικής μορφής

$$\mathbf{d} = \mathbf{g}(\mathbf{m})$$

με δεδομένα και μετρήσεις ασυσχέτιστα

Μία επαναληπτική μέθοδος

Περίπτωση άμεσης μη γραμμικής μορφής

$$\mathbf{d} = \mathbf{g}(\mathbf{m})$$

με δεδομένα και μετρήσεις ασυσχέτιστα

$$\mathbf{m}_{n+1}^{est} = \langle \mathbf{m} \rangle + \mathbf{G}_n^{-g} \left\{ \mathbf{d} - \mathbf{g}(\mathbf{m}_n^{est}) + \mathbf{G}_n [\mathbf{m}_n^{est} - \langle \mathbf{m} \rangle] \right\}$$

$$\begin{aligned} \mathbf{G}_n^{-g} &= [\text{cov } \mathbf{m}] \mathbf{G}_n^T \left\{ [\text{cov } \mathbf{d}] + \mathbf{G}_n [\text{cov } \mathbf{m}] \mathbf{G}_n^T \right\}^{-1} = \\ &= \left\{ [\text{cov } \mathbf{m}]^{-1} + \mathbf{G}_n^T [\text{cov } \mathbf{d}]^{-1} \mathbf{G}_n \right\}^{-1} \mathbf{G}_n^T [\text{cov } \mathbf{d}]^{-1} \end{aligned}$$

$$[\mathbf{G}_n]_{ij} = \partial g_i / \partial m_j$$

Υπολογίζεται στο \mathbf{m}_n^{est}

Μία επαναληπτική μέθοδος

Για μη ακριβή θεωρία

$$[\text{cov } \mathbf{d}] \longrightarrow [\text{cov } \mathbf{d}] + [\text{cov } \mathbf{g}]$$

Για όχι ισχυρή μη γραμμικότητα :

$$[\text{cov } \mathbf{m}_n^{est}] \approx \mathbf{G}_n^{-g} [\text{cov } \mathbf{d}] \mathbf{G}_n^{-gT} + [\mathbf{I} - \mathbf{R}_n] [\text{cov } \mathbf{m}] [\mathbf{I} - \mathbf{R}_n]^T$$

$$\mathbf{R}_n = \mathbf{G}_n^{-g} \mathbf{G}_n$$

Οι πίνακες ανάλυσης δεδομένων και παραμέτρων δεν έχουν εδώ την ίδια σημασία με το γραμμικό πρόβλημα

Μία επαναληπτική μέθοδος

Μια άλλη οπτική της εκτίμησης της λύσης ενός αντίστροφου προβλήματος με επαναληπτική διαδικασία :

$$\mathbf{d} = \mathbf{g}(\mathbf{m})$$

Ανάπτυγμα Taylor γύρω από το \mathbf{m}_n^{est} και αγνοώντας όρους $[\mathbf{m} - \mathbf{m}_n^{est}]^2$ και ανώτερης τάξης

$$\mathbf{g}(\mathbf{m}) \approx \mathbf{g}(\mathbf{m}_n^{est}) + \nabla \mathbf{g}[\mathbf{m} - \mathbf{m}_n^{est}] = \mathbf{g}(\mathbf{m}_n^{est}) + \mathbf{G}_n[\mathbf{m} - \mathbf{m}_n^{est}]$$

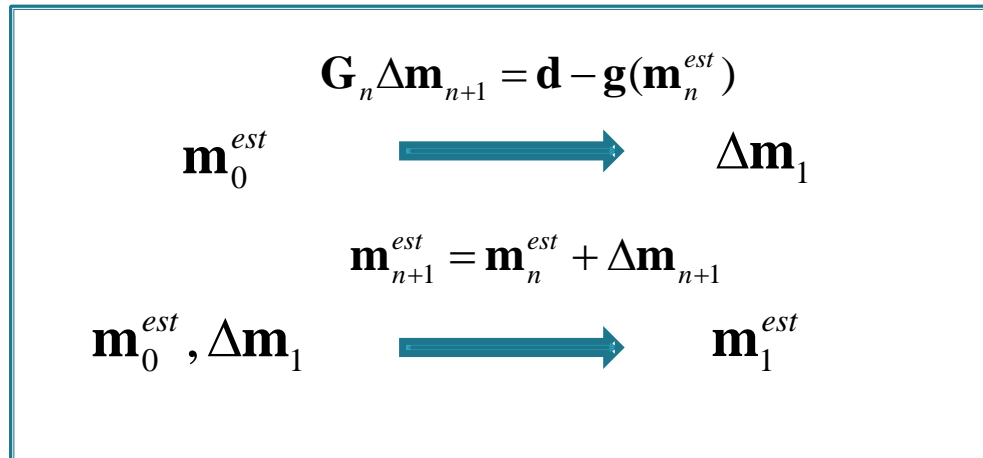
Ορίζοντας $\Delta \mathbf{m}_{n+1} = [\mathbf{m} - \mathbf{m}_n^{est}]$

$$\mathbf{G}_n \Delta \mathbf{m}_{n+1} = \mathbf{d} - \mathbf{g}(\mathbf{m}_n^{est})$$

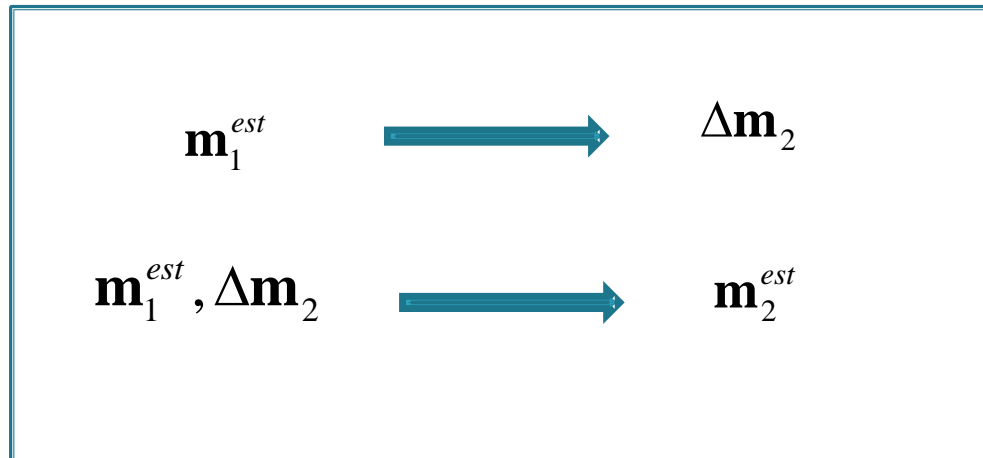
$$\mathbf{m}_{n+1}^{est} = \mathbf{m}_n^{est} + \Delta \mathbf{m}_{n+1}$$

Μία επαναληπτική μέθοδος

$n=0$



$n=1$



Μία επαναληπτική μέθοδος

