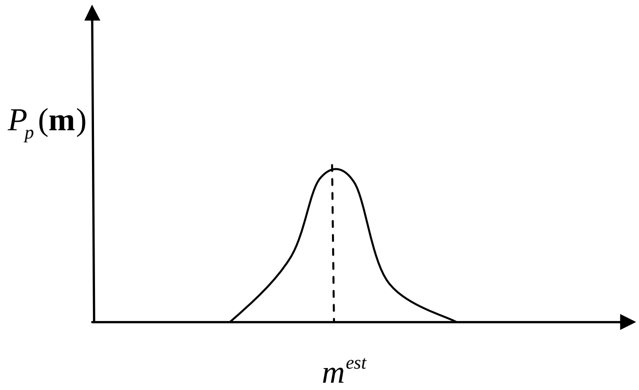
$$P_p(\mathbf{m}) = \int P_T(\mathbf{m},\mathbf{d})\partial\mathbf{d} \tag{5.14}$$

Σε περίπτωση ακριβούς θεωρίας, η εκτίμηση μέγιστης πιθανοφάνειας που δίδει η 5.14 δεν είναι διαφορετική από αυτή που δίδεται εάν μεγιστοποιήσομε την 5.12 στην επιφάνεια που ορίζει η ακριβής θεωρία.



Σχήμα 5.3 Η κατανομή $P_p(\vec{m})$. Το μέγιστο της κατανομής ορίζει την εκτίμηση μέγιστης πιθανοφάνειας m^{est}.

5.6 Η απλή περίπτωση με κανονική κατανομή και γραμμική θεωρία

Υποθέτομε ότι δεν υπάρχει κάποια πληροφορία για την $P_A(\mathbf{m})$ αλλά ότι τα δεδομένα ακολουθούν κανονική κατανομή, που σημαίνει :

$$P_A(\mathbf{d}) \propto \exp[-\tfrac{1}{2}(\mathbf{d}-\mathbf{d}^{obs})^T[\operatorname{cov}\mathbf{d}]^{-1}(\mathbf{d}-\mathbf{d}^{obs})]. \tag{5.15}$$

Εάν δεν υπάρχει λάθος στο μοντέλο που έχει επιλεγεί και για το μοντέλο ισχύσει η σχέση $\mathbf{Gm = d}$, η στατιστική κατανομή εκφράζεται με συνάρτηση δέλτα της μορφής :

$$P_g(\mathbf{m}|\mathbf{d}) = \delta[\mathbf{Gm - d}] \tag{5.16}$$

Η συνάρτηση πιθανοφάνειας θα είναι τότε :

$$P_T(\mathbf{m},\mathbf{d}) \propto \exp[-\tfrac{1}{2}(\mathbf{d}-\mathbf{d}^{obs})^T[\operatorname{cov}\mathbf{d}]^{-1}(\mathbf{d}-\mathbf{d}^{obs})]\delta[\mathbf{Gm - d}] \tag{5.17}$$

Ολοκληρώνοντας ως προς $\mathbf{d}$ παίρνομε από τις ιδιότητες της συνάρτησης δέλτα :

$$P_T(\mathbf{m}) \propto \exp[-\tfrac{1}{2}(\mathbf{Gm}-\mathbf{d}^{obs})^T[\operatorname{cov}\mathbf{d}]^{-1}(\mathbf{Gm}-\mathbf{d}^{obs})] \tag{5.18}$$

Η μεγιστοποίηση της συνάρτησης πιθανοφάνειας συνεπώς στην περίπτωση αυτή, αντιστοιχεί σε λύση ελαχίστων τετραγώνων.

5.7 Η Γενική γραμμική περίπτωση με κανονικές κατανομές.

Υποθέτομε ότι για όλες τις κατανομές ισχύουν οι σχέσεις :

$$P_A(\mathbf{m}) \propto \exp[-\tfrac{1}{2}(\mathbf{m} - <\mathbf{m}>)^T [\mathrm{cov}\,\mathbf{m}]^{-1}(\mathbf{m} - <\mathbf{m}>)] \tag{5.19}$$

$$P_A(\mathbf{d}) \propto \exp[-\tfrac{1}{2}(\mathbf{d} - \mathbf{d}^{obs})^T [\mathrm{cov}\,\mathbf{d}]^{-1}(\mathbf{d} - \mathbf{d}^{obs})] \tag{5.20}$$

$$P_g(\mathbf{m}|\mathbf{d}) \propto \exp[-\tfrac{1}{2}(\mathbf{d} - \mathbf{Gm})^T [\mathrm{cov}\,\mathbf{g}]^{-1}(\mathbf{d} - \mathbf{Gm})] \tag{5.21}$$

Μπορούμε να αποδείξουμε ότι η συνδυασμένη κατανομή που δίνει και τη συνάρτηση πιθανοφάνειας είναι επίσης κανονική :

$$P_T(\mathbf{m}) = P_g(\mathbf{m}|\mathbf{d}) P_A(\mathbf{m}) P_A(\mathbf{d}) \tag{5.22}$$

Με βάση την αρχή της μέγιστης πιθανοφάνειας, μπορούμε τώρα να δώσουμε λύση για τις εκτιμήσεις των παραμέτρων από τη σχέση :

$$\mathbf{m}^{est} = \langle \mathbf{m} \rangle + \mathbf{G}^{-g}[\mathbf{d}^{obs} - \mathbf{G}\langle \mathbf{m} \rangle] = \mathbf{G}^{-g}\mathbf{d} + [\mathbf{I} - \mathbf{R}]\langle \mathbf{m} \rangle \tag{5.23}$$

όπου

$$\begin{aligned} \mathbf{G}^{-g} &= [\mathrm{cov}\,\mathbf{m}]\mathbf{G}^T \left\{ [\mathrm{cov}\,\mathbf{d}] + [\mathrm{cov}\,\mathbf{g}] + \mathbf{G}[\mathrm{cov}\,\mathbf{m}]\mathbf{G}^T \right\}^{-1} \\ &= \left\{ \mathbf{G}^T [[\mathrm{cov}\,\mathbf{d}] + [\mathrm{cov}\,\mathbf{g}]]^{-1}\mathbf{G} + [\mathrm{cov}\,\mathbf{m}]^{-1} \right\}^{-1} \mathbf{G}^T \left\{ [\mathrm{cov}\,\mathbf{d}] + [\mathrm{cov}\,\mathbf{g}] \right\}^{-1} \end{aligned} \tag{5.24}$$

Σημειώνουμε ότι οι δύο μορφές του γενικευμένου αντίστροφου πίνακα είναι ισοδύναμες.

Ο πίνακας συνδιακύμανσης των παραμέτρων υπολογίζεται από την παρατήρηση ότι οι παράμετροι προκύπτουν ως γραμμικός συνδυασμός παρατηρήσεων και εκ προοιμίου δεδομένων και προκύπτει ως :

$$[\mathrm{cov}\,\mathbf{m}^{est}] = \mathbf{G}^{-g}[\mathrm{cov}\,\mathbf{d}]\mathbf{G}^{-gT} + [\mathbf{I} - \mathbf{R}][\mathrm{cov}\,\mathbf{m}][\mathbf{I} - \mathbf{R}]^T \tag{5.25}$$

που διαφέρει από προηγούμενες αντίστοιχες εκφράσεις (βλ. 3.52) ως προς την ύπαρξη της εκ προοιμίου πληροφορίας για τις προς ανάκτηση παραμέτρους.

5.8 Ειδικές περιπτώσεις

Θα εξετάσομε τώρα τη μορφή που παίρνει η λύση μέγιστης πιθανοφάνειας σε ορισμένες απλές περιπτώσεις :

5.8.1 Ακριβή δεδομένα και ακριβές μοντέλο (θεωρία)

Εάν υποθέσομε ότι ισχύει $\sigma_d^2 = \sigma_g^2 = 0$ στη γενική περίπτωση που έχομε εκ προοιμίου πληροφορία για τις παραμέτρους, παίρνομε λύση της μορφής :

$$\mathbf{m}^{est} = \mathbf{G}^{-g}\mathbf{d}^{obs} + [\mathbf{I} - \mathbf{R}]\langle \mathbf{m} \rangle \tag{5.16}$$

όπου η έκφραση του γενικευμένου αντιστρόφου εξαρτάται από το είδος του προβλήματος (υπερ-ορισμένο ή υπο-ορισμένο) κατά τα γνωστά. Με την ίδια λογική, ο δεύτερος όρος στην έκφραση του 5.16 είναι 0 όταν έχομε ένα υπερ-ορισμένο πρόβλημα με ακριβείς μετρήσεις που σημαίνει ότι η λύση μέγιστης πιθανοφάνειας δεν λαμβάνει υπ' όψιν της την αρχική πληροφορία για τις παραμέτρους. Φυσικά στην περίπτωση του υπο-ορισμενου προβλήματος η πληροφορία αυτή συμμετέχει στη λύση.

5.8.2 Μη ακριβές μοντέλο και μη ακριβή δεδομένα.

Μια άλλη ακραία περίπτωση (μη ρεαλιστική ωστόσο σε πρακτικές εφαρμογές) είναι να έχομε τελείως ανακριβές μοντέλο και τελείως ανακριβή δεδομένα. Τότε βέβαια θα ισχύει $\sigma_d^2 \to \infty, \sigma_g^2 \to \infty$ και η λύση μέγιστης πιθανοφάνειας προέρχεται αποκλειστικά από την αρχική πληροφορία :

$$\mathbf{m}^{est} = \langle \mathbf{m} \rangle \tag{5.17}$$

5.8.3 Μη ακριβής αρχική πληροφορία για τις παραμέτρους.

Μπορεί να έχομε και μια άλλη ακραία περίπτωση σύμφωνα με την οποία δεν εμπιστευόμαστε την αρχική πληροφορία για τις παραμέτρους. Αυτό είναι ισοδύναμο με το να θεωρήσουμε ότι ισχύσει μια σχέση της μορφής $\sigma_m^2 \to \infty$.

Όπως μπορεί να δειχθεί στην περίπτωση αυτή παίρνομε λύσεις μέγιστης πιθανοφάνειας της μορφής :

$$\mathbf{m}^{est} = \mathbf{G}^T[\mathbf{G}\mathbf{G}^T]^{-1}\mathbf{d}^{obs} + \{\mathbf{I} - \mathbf{G}^T[\mathbf{G}\mathbf{G}^T]^{-1}\mathbf{G}\}\langle \mathbf{m} \rangle \tag{5.18}$$

που αντιστοιχεί σε λύση ελαχίστου μήκους και έχει τη μορφή της 5.16. Μπορεί να δειχτεί ότι για υπερορισμένο πρόβλημα με τις ίδιες υποθέσεις παίρνομε πάλι λύση της μορφής 5.16 με αλλαγή της μορφής του γενικευμένου αντίστροφου πίνακα.