Α. Δερμάνης

Συνορθώσεις Παρατηρήσεων και Θεωρία Εκτίμησης

τόμος 1

εκδόσεις Ζήτη

ΘΕΣΣΑΛΟΝΙΚΗ

Πρόλογος

Η τοπογραφία, η γεωδαισία, η φωτογραμμετρία και η χαρτογραφία είναι επιστήμες με έντονα εφαρμοσμένο χαρακτήρα, που στηρίζονται στην ανάλυση δεδομένων που προκύπτουν από παρατηρήσεις. Αντικείμενο του βιβλίου αυτού είναι οι σχετικές μέθοδοι ανάλυσης δεδομένων, καθώς και οι αρχές πάνω στις οποίες στηρίζονται.

Σκοπός των μεθόδων αυτών είναι ο προσδιορισμός εκτιμήσεων για τις σχετικές παραμέτρους, με βάση τα αποτελέσματα των διαθέσιμων μετρήσεων. Επειδή στις παραμέτρους που μπορούν να εκτιμηθούν συμπεριλαμβάνονται και τα ίδια τα παρατηρούμενα μεγέθη, και οι εκτιμήσεις τους είναι βελτιωμένες σε σχέση με τις τιμές που προκύπτουν απευθείας από τις μετρήσεις, προκύπτει ταυτόχρονα και μια συνόρθωση των παρατηρήσεων.

Οι μέθοδοι συνόρθωσης των παρατηρήσεων και εκτίμησης παραμέτρων ξεκινούν ουσιαστικά από τη «μέθοδο των ελαχίστων τετραγώνων» των Legendre και Gauss στην αρχή του προηγούμενου αιώνα και γνωρίζουν σημαντική εξέλιξη τα τελευταία 20 χρόνια, ενώ στο μεταξύ εμπλουτίζονται από τις δυνατότητες ερμηνείας που έδωσε η ανάπτυξη της μαθηματικής στατιστικής από τις αρχές του αιώνα μας και μετά.

Η ανάπτυξη της ύλης στηρίχθηκε, ουσιαστικά, στην πεποίθηση ότι επιστημονικό έργο δεν είναι η μηχανιστική εφαρμογή συγκεκριμένων μεθόδων σε συγκεκριμένα προβλήματα, αλλά η βαθύτερη κατανόηση των μεθόδων, που κάνει δυνατή την εφαρμογή τους σε οποιοδήποτε μη τυποποιημένο πρόβλημα. Έτσι, παράλληλα με την παράθεση των σχετικών υπολογιστικών αλγορίθμων, έγινε συστηματική προσπάθεια για την κατανόηση των αρχών πάνω στις οποίες στηρίζονται οι αλγόριθμοι αυτοί και για την ερμηνεία των αποτελεσμάτων που προκύπτουν από αυτούς.

Ο 1ος τόμος ασχολείται με τις αρχές ανάλυσης δεδομένων, με βάση το απλούστερο δυνατό πρόβλημα της εκτίμησης ενός μεγέθους από επαναλαμβανόμενες μετρήσεις (κεφάλαια 2 και 5). Παράλληλα, δίνονται οι βασικές έννοιες της μαθηματικής στατιστικής (κεφάλαιο 3), που είναι απαραίτητες τόσο για την κατανόηση των αρχών πάνω στις οποίες στηρίζονται οι μέθοδοι ανάλυσης δεδομένων, όσο και για την ερμηνεία των αποτελεσμάτων τους (κεφάλαιο 6). Στην εισαγωγή (κεφάλαιο 1) γίνεται μια απόπειρα να τοποθε-

τηθούν οι παραπάνω μέθοδοι στα πλαίσια των γενικότερων μεθόδων και των προβλημάτων των εφαρμοσμένων επιστημών και της επιστήμης γενικότερα.

Στο 2ο τόμο δίνονται αρχικά οι βασικές μέθοδοι συνόρθωσης των παρατηρήσεων (κεφάλαια 7 έως 11) και αναλύεται ο χαρακτήρας τους από φυσική, αλγεβρική, στατιστική και γεωμετρική σκοπιά (κεφάλαια 12, 13 και 15). Επίσης, δίνονται οι μέθοδοι ερμηνείας και αξιολόγησης των αποτελεσμάτων της συνόρθωσης (κεφάλαιο 14), με βάση τις αρχές που ήδη αναπτύχθηκαν στον 1ο τόμο (κεφάλαιο 6).

Τέλος, ξεκινώντας από το καθαρά ντετερμινιστικό πρόβλημα της παρεμβολής, δίνονται οι πιο σύγχρονες μέθοδοι συνόρθωσης και πρόγνωσης με στοχαστικές παραμέτρους (κεφάλαια 16 και 17), που αποκτούν συνεχώς πλατύτερη εφαρμογή ακόμα και σε απλά τοπογραφικά προβλήματα, και που είναι γνωστές με το γενικό όρο «σημειακή προσαρμογή».

Το βιβλίο αυτό προέκυψε από την παραπέρα ανάπτυξη και τον εμπλουτισμό των σημειώσεων που από το 1982 δίνονται στους φοιτητές του Τμήματος Αγρονόμων και Τοπογράφων Μηχανικών για τις ανάγκες των μαθημάτων «Αρχές Ανάλυσης Δεδομένων» και «Συνορθώσεις Παρατηρήσεων και θεωρία Εκτίμησης».

Ευχαριστίες οφείλονται στη Βάσω Δερμάνη για τις προσπάθειές της να βελτιώσει γλωσσικά το κείμενο και στο Δημήτρη Ρωσσικόπουλο για την κάθε είδους βοήθειά του στην προετοιμασία του βιβλίου.

Ελπίζουμε ότι μέσα από την κριτική και τις υποδείξεις των αναγνωστών και των συναδέλφων, θα μας δοθεί στο μέλλον η δυνατότητα για μια νέα έκδοση με βελτίωση τόσο του ουσιαστικού μέρους όσο και της αναγνωσιμότητας του βιβλίου αυτού.

Θεσσαλονίκη, Μάης 1986

Α. Δερμάνης

ΠΕΡΙΕΧΟΜΕΝΑ

Συνοπτικοί πίνακες

1

Εισαγωγή

Αντικείμενο του βιβλίου αυτού είναι η μελέτη των μεθόδων που χρησιμοποιούνται, κυρίως στην τοπογραφία, αλλά και ευρύτερα στις γεωδαιτικές επιστήμες (γεωδαισία, φωτογραμμετρία, χαρτογραφία), για την ανάλυση των δεδομένων που προκύπτουν από μετρήσεις. Πριν όμως από τη μελέτη συγκεκριμένων μεθόδων, θα εξετάσουμε τις γενικότερες αρχές ανάλυσης δεδομένων.

Οι μέθοδοι ανάλυσης δεδομένων δεν αρκεί να είναι γνωστές απλά σαν τεχνικές υπολογισμών, αλλά είναι απαραίτητη και μια βαθύτερη κατανόηση των γενικότερων αρχών πάνω στις οποίες στηρίζονται, γιατί τότε μόνο είναι δυνατή η ερμηνεία και η σωστή χρήση των αποτελεσμάτων που προκύπτουν από τους υπολογισμούς. Η ερμηνεία αυτή στηρίζεται στην κατανόηση του χαρακτήρα, των δυνατοτήτων, αλλά και των περιορισμών της κάθε συγκεκριμένης μεθόδου.

Οι μέθοδοι ανάλυσης δεδομένων που χρησιμοποιούνται στην τοπογραφία και τις συγγενείς επιστήμες, παρά μια φαινομενική ομοιότητα με μεθόδους που χρησιμοποιούνται στην οικονομετρία και την βιομετρία, έχουν μια ιδιαιτερότητα, που οφείλεται στο γεγονός ότι αναπτύχθηκαν ιστορικά μαζί με τις επιστήμες αυτές. Δεν έχουμε δηλαδή να κάνουμε με τη συνηθισμένη μεταφορά έτοιμων μεθόδων της στατιστικής σε συγκεκριμένες εφαρμογές, όπως γίνεται σε τόσες άλλες εφαρμοσμένες επιστήμες. Κύριο χαρακτηριστικό των μεθόδων αυτών είναι ο ποσοτικός προσδιορισμός ορισμένων φυσικών μεγεθών, κυρίως γεωμετρικών, με τη με-

μεγαλύτερη δυνατή ακρίβεια. Την ιδιαιτερότητα αυτή μαρτυρεί και ο όρος «συνόρθωση των παρατηρήσεων», που είναι όρος καθαρά τοπογραφικός-γεωδαιτικός.

Από την άλλη μεριά, παρά τη σημερινή διαίρεση της επιστήμης σε διάφορους κλάδους, οι μέθοδοι που χρησιμοποιούν οι επιμέρους επιστήμες δεν είναι ξεκομμένες από τα γενικότερα κοινά προβλήματα των εφαρμοσμένων επιστημών, αλλά και της επιστήμης σαν σύνολο. Ούτε όμως και η επιστήμη, σαν ιδιαίτερη ανθρώπινη δραστηριότητα, είναι ξεκομμένη από τις υπόλοιπες ανθρώπινες δραστηριότητες, όπως αυτές καθορίζονται μέσα στην κάθε φορά συγκεκριμένη ιστορική (κοινωνική, πολιτική, οικονομική, πολιτιστική) πραγματικότητα.

Τέλος, επειδή η επιστήμη έχει να κάνει με τη γνώση, δεν μπορεί κανείς να αποφύγει εντελώς κάποια ερωτήματα σχετικά με τη δυνατότητα, σε τελευταία ανάλυση, του ανθρώπου για γνώση, αγγίζοντας έτσι προβλήματα της φιλοσοφίας και ειδικότερα της φιλοσοφίας της επιστήμης.

Για τους παραπάνω λόγους, θα προσπαθήσουμε, στο εισαγωγικό αυτό κεφάλαιο, να θέσουμε μάλλον ερωτήματα παρά να δώσουμε απαντήσεις – κάτι τέτοιο θα ήταν εξάλλου υπέρμετρα φιλόδοξο. Θα σκιαγραφήσουμε έτσι, με χοντρές πινελιές, τη μεγάλη σύνθεση του επιστημονικού στοχασμού και της επιστημονικής πράξης, όπου μια μικρή μόνο λεπτομέρεια, αλλά για μας τουλάχιστον σημαντική, είναι οι αρχές και οι μέθοδοι ανάλυσης δεδομένων. Τα ερωτήματα με τα οποία θα ασχοληθούμε είναι: Τί είναι η επιστήμη σήμερα, ποιο είναι το αντικείμενο και ποιες οι μέθοδοί της, και πως απόκτησε, μέσα από την ιστορική εξέλιξη, το σημερινό της περιεχόμενο. Ποιες είναι, ειδικότερα, οι μέθοδοι των εφαρμοσμένων επιστημών, και ποια η σχέση των μεθόδων που χρησιμοποιούνται στην τοπογραφία και στις γεωδαιτικές επιστήμες με τις μεθόδους αυτές.

1.1. Το περιεχόμενο της επιστήμης

Το ερώτημα τί είναι επιστήμη είναι ένα ερώτημα που φέρνει σε αμηχανία κάθε επιστήμονα. Τόσο μάλιστα μεγαλύτερη είναι η αμηχανία και η αδυναμία διατύπωσης μιας συγκεκριμένης απάντησης, όσο μεγαλύτερος και εντονότερος είναι ο προβληματισμός του επιστήμονα πάνω στο αντικείμενο της δουλειάς του. Κάθε ορισμός της επιστήμης, όντας αναγκαστικά συνοπτικός, δεν μπορεί να είναι απόλυτα σαφής. Ένας ορισμός, αν και περικλείνει τον κίνδυνο παρανόησης, είναι χρήσιμος σαν αφετηρία. Θα παραθέσουμε λοιπόν μερικούς ορισμούς από τη βιβλιογραφία:

> «Σύνολο των ανθρωπίνων γνώσεων για τη φύση, την κοινωνία και τη σκέψη, που αποκτιέται με την ανακάλυψη των αντικειμενικών νόμων των φαινομένων και την εξήγησή τους».
> (Λεξικό Petit Larousse, λήμμα «science»)
>
> «Ως επιστήμη εννοούμε τη διανοητική εκείνη δραστηριότητα με την οποία οι άνθρωποι επιχειρούν μια συστηματική περιγραφή, ταξινόμηση και ερμηνεία παρατηρημένων συμβάντων».
> (Βέικος, 1982, σελ. 92)
>
> «Επιστήμη είναι μια κοινωνική πρακτική που αποβλέπει στο να ερευνήσει και να διατυπώσει τους νόμους μιας ειδικής περιοχής της πραγματικότητας και, ταυτόχρονα, το απόκτημα αυτής της πρακτικής».
> (Μπιτσάκης, 1983, σελ. 69).

Η επιστήμη έχει να κάνει λοιπόν με την απόκτηση γνώσης, αλλά κάθε μορφή ανθρώπινης γνώσης δεν είναι και επιστήμη. Η επιστήμη ασχολείται με τη συγκροτημένη γνώση που δεν περιγράφει απλά τα φυσικά φαινόμενα, αλλά προσπαθεί και να τα «εξηγήσει».

Δύο ακόμη καθοριστικά στοιχεία της επιστήμης είναι οι μέθοδοι και οι σκοποί της. Οι μέθοδοι όλων των επιστημών δεν είναι ενιαίες, αλλά η καθεμιά απ' αυτές διαμορφώνει τις δικές της μεθόδους, που οπωσδήποτε όμως έχουν κάποια κοινά χαρακτηριστικά. Οι κύριες διαφορές εμφανίζονται ανάμεσα στις δύο ομάδες, στις οποίες μπορούν να διακριθούν οι επιμέρους επιστήμες: στις φυσικές επιστήμες και στις επιστήμες του ανθρώπου. Οι φυσικές επιστήμες καταγίνονται με τα φυσικά φαινόμενα της ανόργανης ύλης και των έμβιων οργανισμών, ενώ οι επιστήμες του ανθρώπου μελετούν τα φαινόμενα που έχουν σχέση με τις δραστηριότητες των ανθρώπων σε οργανωμένα κοινωνικά σύνολα. Διαφορές στις μεθόδους εμφανίζονται επίσης ανάμεσα στις θεωρητικές επιστήμες και στις εφαρμοσμένες επιστήμες. Οι θεωρητικές επιστήμες ασχολούνται με την «εξήγηση» των φαινομένων μέσα από τη διατύπωση νόμων. Οι εφαρμοσμένες επιστήμες ασχολούνται με την αξιοποίηση των επιστημονικών γνώσεων για την ικανοποίηση των ανθρώπινων αναγκών.

Σχετικά με τους σκοπούς της επιστήμης υπάρχουν διάφορες απόψεις, που βρίσκονται λιγότερο ή περισσότερο κοντά σε μια από δύο εκ διαμέτρου αντίθετες θέσεις.

Η πρώτη βλέπει σαν κίνητρο της επιστημονικής πρακτικής την έμφυτη επιθυμία του ανθρώπου για γνώση (η γνώση για τη γνώση με άλλα λόγια)· επιθυμία που είναι ανεξάρτητη από το αν τελικά η γνώση που αποκτιέται χρησιμοποιείται και για την ικανοποίηση υλικών αναγκών.

Η δεύτερη βλέπει σαν κινητήρια δύναμη την ικανοποίηση των υλικών αναγκών, που οδηγεί τον άνθρωπο στην αναζήτηση γνώσεων και στην έρευνα και συνεισφέρει τελικά στην «καθαρή» γνώση.

Η αλήθεια πρέπει να βρίσκεται κάπου ανάμεσα. Ο επιστήμονας, σαν άτομο, οδηγείται σίγουρα από την προσωπική ικανοποίηση που του δίνει η επιστημονική ανακάλυψη, είτε άμεσα είτε και μέσα από την επακόλουθη επιστημονική αναγνώριση, άσχετα με το αν η δραστηριότητα αυτή αποτελεί και μέσο βιοπορισμού. Η εικόνα ενός επιστήμονα που εργάζεται για λόγους μόνο βιοποριστικούς, όχι μόνο δεν είναι ελκυστική, αλλά δεν είναι ούτε ρεαλιστική. Όμως, η επιλογή του είδους της έρευνας, σχεδόν ποτέ δεν αποτελεί επιλογή του ίδιου του επιστήμονα, κάτι που είναι περισσότερο αληθινό σήμερα παρά ποτέ. Ακόμη και όταν ο ίδιος πιστεύει ότι κάνει μια ελεύθερη επιλογή, η επιλογή αυτή δεν είναι ανεξάρτητη από τις κατευθύνσεις που του υποβάλλονται από τα γενικότερα προβλήματα της επιστήμης του στην συγκεκριμένη περίοδο, όπως διαμορφώνονται από τις μέχρι στιγμής γνώσεις και τις σχετικές έρευνες των συναδέλφων του.

Έτσι, αν ξεφύγουμε από την κλίμακα του μεμονωμένου επιστήμονα, οι σκοποί και ο ρόλος, γενικότερα, της επιστημονικής έρευνας διαμορφώνονται μέσα στα πλαίσια της συγκεκριμένης ιστορικής πραγματικότητας. Αναμφισβήτητα, το σημαντικότερο μέρος της επιστημονικής έρευνας σήμερα, ακόμη και σε επιστημονικές περιοχές όπου κάτι τέτοιο δύσκολα γίνεται φανερό (όπως, π.χ., στην ψυχολογία), ακολουθεί τις κατευθύνσεις που επιβάλλουν οι ανάγκες της βιομηχανικής παραγωγής και, σε μεγάλο βαθμό δυστυχώς, οι παράλογες απαιτήσεις των πολεμικών εξοπλισμών. Έτσι εξάλλου εξηγείται και ως ένα βαθμό το γιατί η επιστήμη, όπως την εννοούμε σήμερα, δεν αναπτύχθηκε και σε παλιότερους πολιτισμούς, παρά τα αξιόλογα επιτεύγματά τους.

Σε τελευταία ανάλυση, η παραπάνω αντιδικία γύρω από τα κίνητρα της επιστημονικής δραστηριότητας, δεν είναι παρά αντανάκλαση μιας βαθύτερης πάλης ανάμεσα στα δύο αντίπαλα φιλοσοφικά ρεύματα, τον ιδεαλισμό και τον υλισμό.

2

Κριτήρια βέλτιστης εκτίμησης παραμέτρων

2.1. Γενικά

Το πρόβλημα του προσδιορισμού των τιμών άγνωστων παραμέτρων, που περιγράφουν κάποιο φυσικό σύστημα χρησιμοποιώντας έναν αριθμό παρατηρήσεων μεγαλύτερο από τον αριθμό των παραμέτρων, δεν έχει στη γενική περίπτωση καμιά λύση. Είναι δηλαδή αδύνατο να βρεθούν τιμές για τις άγνωστες παραμέτρους τέτοιες, ώστε υπολογίζοντας από αυτές τις παρατηρημένες ποσότητες να προκύπτουν τιμές ίσες με τα εξαγόμενα των παρατηρήσεων. Αντίθετα, εξαιτίας των σφαλμάτων που αναπόφευκτα υπεισέρχονται στις παρατηρήσεις, παρουσιάζονται αποκλίσεις ανάμεσα στις παρατηρημένες τιμές και αυτές που κάθε φορά υπολογίζονται με βάση κάποια επιλογή τιμών για τις άγνωστες παραμέτρους.

Κάθε επιλογή τιμών για τις άγνωστες παραμέτρους διαφέρει από τις αληθινές τιμές τους, που, εξαιτίας των σφαλμάτων των παρατηρήσεων, παραμένουν απροσδιόριστες. Είναι μόνο **εκτιμήσεις** των αληθινών τιμών, δηλαδή τιμές που προκύπτουν κατά εκτίμηση. Μία τέτοια εκτίμηση δεν πρέπει φυσικά να είναι αυθαίρετη, αλλά να γίνεται έτσι ώστε να πλησιάζει τις άγνωστες αληθινές τιμές.

Τα σφάλματα των παρατηρήσεων είναι γενικά μικρές σε απόλυτη τιμή ποσότητες, επειδή ακριβώς καταβάλλονται προσπάθειες να μειωθεί όσο το δυνατόν περισσότερο η επίδρασή τους, τόσο κατά την κατασκευή των μετρητικών οργάνων, όσο και κατά το σχεδιασμό και την εκτέλεση των παρατηρήσεων. Οι διαφορές ανάμεσα στις παρατηρημένες τιμές και στις τιμές που υπολογίζονται από τις κατά εκτίμηση τιμές των παραμέτρων, αποτελούν εκτιμήσεις των άγνωστων αληθινών τιμών των σφαλμάτων, γιαυτό και πρέπει να είναι μικρές σε απόλυτη τιμή.

Αν εισαχθεί κάποια ποσότητα που να μετρά το πόσο μικρές είναι οι εκτιμήσεις των σφαλμάτων, θα πρέπει να αναζητηθούν εκείνες οι τιμές των παραμέτρων που κάνουν την ποσότητα αυτή όσο το δυνατόν μικρότερη. Η ελαχιστοποίηση μιας τέτοιας ποσότητας ονομάζεται **κριτήριο βέλτιστης εκτίμησης** και οδηγεί στις **βέλτιστες τιμές** των παραμέτρων. Η όλη διαδικασία προσδιορισμού των βέλτιστων τιμών των παραμέτρων, με βάση το κριτήριο βέλτιστης εκτίμησης, ονομάζεται **βέλτιστη εκτίμηση** των παραμέτρων.

Μια ιδιότητα, απαραίτητη για κάθε προτεινόμενο κριτήριο βέλτιστης εκτίμησης, είναι το να οδηγεί σε ένα μοναδικά καθορισμένο σύνολο βέλτιστων τιμών των παραμέτρων. Να μην υπάρχουν δηλαδή δύο ή περισσότερα διαφορετικά σύνολα τιμών των παραμέτρων που να δίνουν την ίδια ελάχιστη τιμή, ικανοποιώντας ταυτόχρονα το κριτήριο βελτιστοποίησης.

Με τον προσδιορισμό των βέλτιστων τιμών των παραμέτρων μπορούν να υπολογιστούν και οι αντίστοιχες βέλτιστες τιμές των παρατηρημένων μεγεθών. Έτσι, οι τιμές που προέκυψαν από τις παρατηρήσεις αντικαθίστανται από τις βέλτιστες κατά εκτίμηση τιμές τους και η διαδικασία αυτή ονομάζεται **συνόρθωση των παρατηρήσεων**. Η έννοια της συνόρθωσης των παρατηρήσεων είναι, κατά κάποιο τρόπο, ευρύτερη από την έννοια της βέλτιστης εκτίμησης των παραμέτρων. Όπως θα δούμε παρακάτω, οι παρατηρήσεις μπορούν να συνορθωθούν και χωρίς να καταφύγουμε σε παραμέτρους που να περιγράφουν το φυσικό σύστημα που μελετάμε.

Για να εξετάσουμε και να αξιολογήσουμε διάφορα κριτήρια βελτιστοποίησης, θα ασχοληθούμε με την απλούστερη δυνατή περίπτωση εκτίμησης παραμέτρων, κεντρώνοντας το ενδιαφέρον μας στην ουσία του προβλήματος, χωρίς τις δυσκολίες που μπορεί να προκύψουν από τις ιδιαιτερότητες κάποιων περισσότερο πολύπλοκων προβλημάτων εκτίμησης παραμέτρων.

Το πιο απλό πρόβλημα εκτίμησης παραμέτρων που μπορούμε να εξετάσουμε, είναι εκείνο όπου κάποια άγνωστη παράμετρος με αληθινή τιμή x^{α} είναι η παρατηρούμενη ποσότητα σε μια σειρά n παρατηρήσεων με εξαγόμενα b_1, b_2, ..., b_n. Για να συγκεκριμενοποιήσουμε το πρόβλημα, μπορούμε να φανταστούμε πως η απόσταση ανάμεσα σε δύο σημεία μετριέται n φορές με τη βοήθεια κάποιου οργάνου μέτρησης αποστάσεων, π.χ. μια μετροταινία.

Το εξαγόμενο κάθε μέτρησης b_i διαφέρει από την αντίστοιχη αληθινή τιμή x^α, εξαιτίας του αληθινού σφάλματος της παρατήρησης v_i^α, σύμφωνα με τη σχέση

$$(1) \qquad b_i = x^\alpha + v_i^\alpha, \qquad i = 1, 2, \ldots, n .$$

Για κάθε τιμή x που επιλέγεται σαν εκτίμηση της αληθινής τιμής x^α, προκύπτουν αντίστοιχα εκτιμήσεις των σφαλμάτων

$$(2) \qquad v_i = b_i - x .$$

Το πρόβλημα είναι να προσδιοριστεί η βέλτιστη τιμή $\hat{x}$ που οδηγεί σε βέλτιστη εκτίμηση τιμών των σφαλμάτων

$$(3) \qquad \hat{v}_i = b_i - \hat{x} ,$$

έτσι ώστε οι απόλυτες τιμές των σφαλμάτων $|\hat{v}_i|$ να είναι όσο το δυνατό μικρότερες, ικανοποιώντας κάποιο εξαρχής επιλεγμένο κριτήριο βελτιστοποίησης.

Θα εξετάσουμε στη συνέχεια τρία κριτήρια βελτιστοποίησης που, χωρίς να είναι τα μοναδικά, είναι τα περισσότερο συνηθισμένα στην πράξη.

2.2. Το κριτήριο των ελαχίστων τετραγώνων

Σύμφωνα με το κριτήριο αυτό, οι κατά εκτίμηση τιμές των σφαλμάτων $\hat{v}_i$ πρέπει να ελαχιστοποιούν την ποσότητα

$$(4) \qquad \varphi = \sum_{i=1}^{n} v_i^2 = v_1^2 + v_2^2 + \ldots + v_n^2$$

σε σχέση με κάθε άλλο παραδεκτό σύνολο τιμών v_i. Το κριτήριο των ελαχίστων τετραγώνων γράφεται συμβολικά

$$(5) \qquad \sum_{i=1}^{n} v_i^2 = \min .$$

Αντικαθιστώντας τα v_i από τη σχέση (2) στη σχέση (4), προκύπτει ότι πρέπει να ελαχιστοποιηθεί η ποσότητα

$$\varphi(x) = \sum_{i=1}^{n} (b_i - x)^2 . \tag{6}$$

Πρέπει δηλαδή να αναζητηθεί η τιμή $\hat{x}$, για την οποία $\varphi(\hat{x}) < \varphi(x)$ για κάθε άλλη τιμή x της άγνωστης παραμέτρου.

Από τη θεωρία των ακρότατων τιμών μιας συνάρτησης είναι γνωστό ότι η συνάρτηση $\varphi(x)$ έχει ελάχιστη τιμή σε σημείο όπου μηδενίζεται η πρώτη παράγωγός της

$$\frac{d\varphi}{dx} = 0 , \tag{7}$$

ενώ η δεύτερη παράγωγός της είναι θετική

$$\frac{d^2\varphi}{dx^2} > 0 . \tag{8}$$

Παραγωγίζοντας τη $\varphi(x)$, όπως δίνεται από τη σχέση (6), προκύπτει ότι

$$\frac{d\varphi}{dx}(x) = 2 \sum_{i=1}^{n} (x - b_i) \tag{9}$$

και

$$\frac{d^2\varphi}{dx^2} = 2n > 0 . \tag{10}$$

Επομένως, η τιμή $\hat{x}$ που ελαχιστοποιεί τη συνάρτηση $\varphi(x)$ δίνεται από τη σχέση

$$\frac{d\varphi}{dx}(\hat{x}) = 2 \sum_{i=1}^{n} (\hat{x} - b_i) = 0 . \tag{11}$$

Λύνοντας ως προς $\hat{x}$ την εξίσωση (11), προκύπτει

$$\hat{x} = \frac{1}{n} \sum_{i=1}^{n} b_i . \tag{12}$$

Σύμφωνα λοιπόν με το κριτήριο των ελαχίστων τετραγώνων, η βέλτιστη τιμή μιας πολλαπλά παρατηρημένης ποσότητας είναι **ο αριθμητικός μέσος όρος** των παρατηρούμενων τιμών.

2.3. Το κριτήριο της ελαχιστοποίησης του μέγιστου (minimax)

Για κάθε επιλογή x της τιμής της άγνωστης παραμέτρου, προκύπτουν αντίστοιχες εκτιμήσεις των τιμών των σφαλμάτων v_i. Κάποιο από τα v_i έχει μέγιστη απόλυτη τιμή που τη συμβολίζουμε με $\max_i |v_i|$. Ένα κριτήριο βέλτιστης εκτίμησης είναι η ελαχιστοποίηση του μέγιστου σφάλματος ή κριτήριο minimax (minimum του maximum), όπως συνήθως ονομάζεται. Το κριτήριο minimax γράφεται συμβολικά

$$\max_i |v_i| = \min . \tag{13}$$

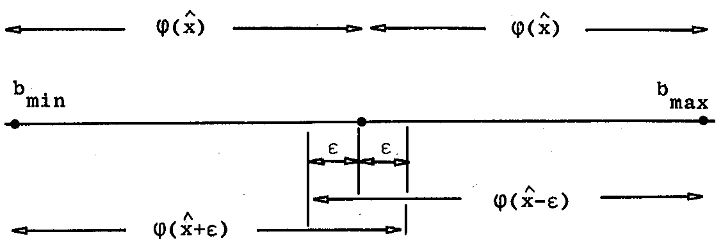
Η βέλτιστη τιμή της παραμέτρου είναι η τιμή $\hat{x}$ για την οποία ελαχιστοποιείται η ποσότητα $\varphi = \max_i |v_i|$, οπότε ικανοποιείται το κριτήριο

$$\varphi(x) = \max_i |b_i - x| = \min . \tag{14}$$

Θα αποδείξουμε ότι,σύμφωνα με το κριτήριο minimax, η βέλτιστη εκτίμηση $\hat{x}$, για το συγκεκριμένο πρόβλημα, δίνεται από τη **διάμεση** τιμή των παρατηρούμενων τιμών

$$\hat{x} = b_{min} + \frac{b_{max} - b_{min}}{2} \tag{15}$$

όπου b_{max} και b_{min} είναι η μέγιστη και η ελάχιστη, αντίστοιχα, από τις παρατηρούμενες τιμές b_i.



Σχήμα 1

Για την τιμή $\hat{x}$ το σφάλμα $\hat{v}_i$ με μέγιστη απόλυτη τιμή είναι

$$\varphi(\hat{x}) = \max_i |\hat{v}_i| = \hat{x} - b_{min} = b_{max} - \hat{x} = \frac{b_{max} - b_{min}}{2} \tag{16}$$

Θα δείξουμε ότι αν η τιμή $\hat{x}$ αυξηθεί ή ελαττωθεί κατά μία μικρή θετική ποσότητα ε, η ποσότητα $\varphi = \max_i |v_i|$ αυξάνεται και στις δύο περιπτώσεις και επομένως ελαχιστοποιείται για την τιμή $x = \hat{x}$. Πράγματι, όπως φαίνεται και στο σχήμα 1

(17) $\varphi(\hat{x}+\varepsilon) = \max_i |v_i| = \hat{x}+\varepsilon-b_{min} > \hat{x}-b_{min} = \varphi(\hat{x})$

(18) $\varphi(\hat{x}-\varepsilon) = \max_i |v_i| = b_{max}-(\hat{x}-\varepsilon) > b_{max}-\hat{x} = \varphi(\hat{x})$.

2.4. Το κριτήριο τής ελαχιστοποίησης του αθροίσματος των απόλυτων τιμών

Το κριτήριο της ελαχιστοποίησης του αθροίσματος των απόλυτων τιμών των σφαλμάτων, ή απλά κριτήριο των ελαχίστων απόλυτων τιμών, γράφεται συμβολικά

$$\sum_{i=1}^{n} |v_i| = \min . \tag{19}$$

Στην περίπτωση αυτή πρέπει να ελαχιστοποιηθεί η ποσότητα

$$\varphi(x) = \sum_{i=1}^{n} |b_i - x| . \tag{20}$$

Η τιμή $\hat{x}$ που ελαχιστοποιεί την ποσότητα $\varphi(x)$, εξαρτάται από το αν ο αριθμός n των παρατηρήσεων είναι άρτιος ή περιττός. Υποθέτουμε πως οι παρατηρήσεις έχουν ταξινομηθεί και αριθμηθεί με αυξανόμενη σειρά μεγέθους, έτσι ώστε

$$b_1 \leq b_2 \leq \ldots \leq b_{n-1} \leq b_n . \tag{21}$$

Στην περίπτωση περιττού αριθμού παρατηρήσεων $n = 2k+1$ η τιμή $\hat{x}$, που ελαχιστοποιεί το άθροισμα των απόλυτων τιμών των αποκλίσεων, είναι η **ενδιάμεση** τιμή παρατήρησης b_{k+1}. Έχουμε δηλαδή

$$\hat{x} = b_{k+1} \tag{22}$$

όπου η τιμή b_{k+1} είναι ενδιάμεση κατά σειρά μεγέθους, αφού υπάρχει ίσος αριθμός k μικρότερων (b_1, b_2, ..., b_k) και k μεγαλύτερων τιμών (b_{k+2}, b_{k+3}, ..., b_{2k+1}) παρατηρήσεων.

Για την τιμή $\hat{x} = b_{k+1}$ η ποσότητα φ έχει την τιμή

$$\varphi(\hat{x}) = \sum_{i=1}^{n} |b_i - b_{k+1}| = \sum_{i=1}^{k} (b_{k+1} - b_i) + \sum_{i=k+2}^{2k+1} (b_i - b_{k+1}) . \tag{23}$$

Θα δείξουμε ότι αν η τιμή του x αυξηθεί ή ελαττωθεί κατά μία μικρή θετική ποσότητα, η αντίστοιχη τιμή της φ αυξάνει και στις δύο περιπτώσεις και κατά συνέπεια η τιμή $\varphi(\hat{x})$ είναι η ελάχιστη δυνατή. Πράγματι

$$(24)\quad \varphi(\hat{x}+\varepsilon) = \sum_{i=1}^{n} |b_i - (b_{k+1}+\varepsilon)| =$$

$$= \sum_{i=1}^{k} (b_{k+1}+\varepsilon-b_i) + [(b_{k+1}+\varepsilon)-b_{k+1}] + \sum_{i=k+2}^{2k+1} (b_i-b_{k+1}-\varepsilon) =$$

$$= \sum_{i=1}^{k} (b_{k+1}-b_i) + k\varepsilon + \varepsilon + \sum_{i=k+2}^{2k+1} (b_i-b_{k+1}) - k\varepsilon =$$

$$= \varphi(\hat{x}) + \varepsilon > \varphi(\hat{x}) .$$

Παρόμοια

$$(25)\quad \varphi(\hat{x}-\varepsilon) = \sum_{i=1}^{n} |b_i - (b_{k+1}-\varepsilon)| =$$

$$= \sum_{i=1}^{k} (b_{k+1}-\varepsilon-b_i) + [b_{k+1}-(b_{k+1}-\varepsilon)] + \sum_{i=k+2}^{2k+1} (b_i-b_{k+1}+\varepsilon) =$$

$$= \sum_{i=1}^{k} (b_{k+1}-b_i) - k\varepsilon + \varepsilon + \sum_{i=k+2}^{2k+1} (b_i-b_{k+1}) + k\varepsilon =$$

$$= \varphi(\hat{x}) + \varepsilon > \varphi(\hat{x}) .$$

Η διαδικασία που ακολουθήθηκε στην παραπάνω απόδειξη γίνεται ευκολότερα κατανοητή με τη βοήθεια του σχήματος 2, όπου οι παρατηρήσεις b_i έχουν τοποθετηθεί πάνω στον άξονα των πραγματικών αριθμών x, ενώ τα οριζόντια ευθύγραμμα τμήματα παριστάνουν τους προσθετέους $|v_i|$ στην ποσότητα φ. Το σύνολο των μηκών των ευθυγράμμων τμημάτων ισούται με την ποσότητα φ, που προφανώς είναι μικρότερη για την τιμή $\hat{x} = b_4$, ενώ αυξάνεται κατά ένα μικρό τμήμα μήκους ε για τις τιμές $\hat{x}+\varepsilon$ ή $\hat{x}-\varepsilon$.

Στην περίπτωση άρτιου αριθμού παρατηρήσεων $n=2k$, η ποσότητα φ ελαχιστοποιείται για οποιαδήποτε τιμή $\hat{x}$ στο διάστημα $b_k \leq \hat{x} \leq b_{k+1}$. Δηλαδή και οι δύο ενδιάμεσες τιμές b_k, b_{k+1}, καθώς και κάθε τιμή ανάμεσά τους, ικανοποιούν τη συνθήκη ελαχιστοποίησης (19). Πρέπει να τονιστεί ότι η αδυναμία της συνθήκης ελαχιστοποίησης να οδηγήσει, στην περίπτωση αυτή, σε μια μοναδικά καθορισμένη τιμή $\hat{x}$ της άγνωστης παραμέτρου, αποτελεί σοβαρό μειονέκτημα.

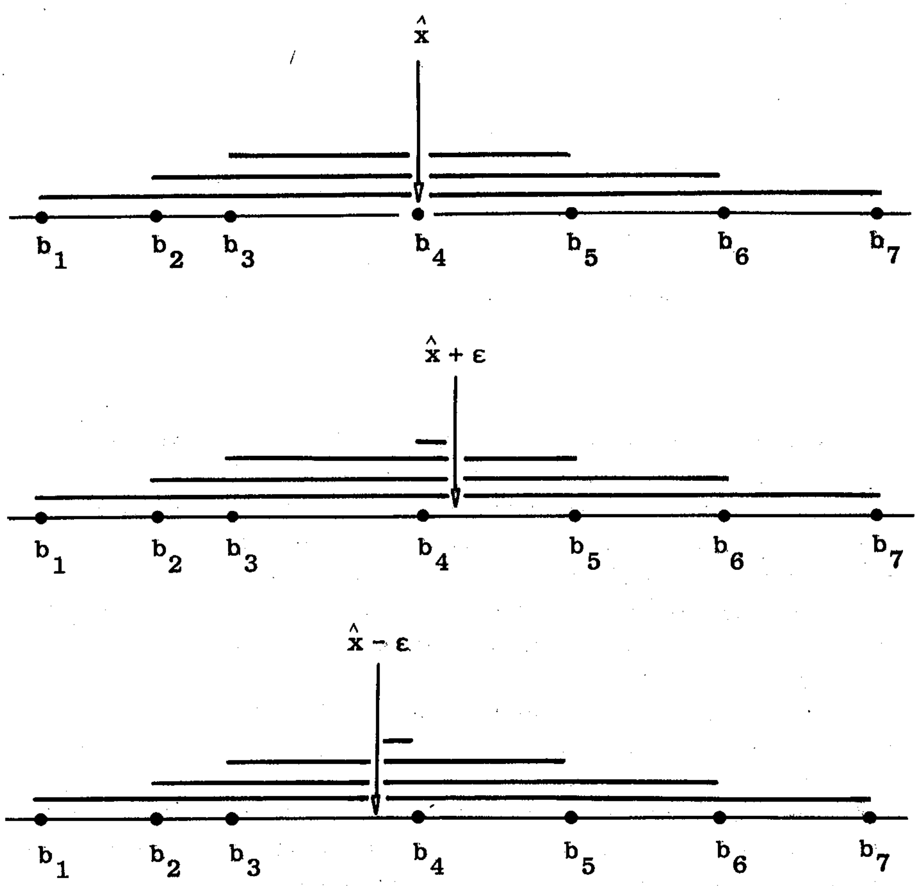
Για οποιαδήποτε τιμή στο διάστημα $b_k \leq \hat{x} \leq b_{k+1}$, έχουμε

$$\varphi(\hat{x}) = \sum_{i=1}^{n} |b_i - \hat{x}| = \sum_{i=1}^{k} (\hat{x} - b_i) + \sum_{i=k+1}^{2k} (b_i - \hat{x}) = \tag{26}$$

$$= \sum_{i=1}^{k} (\hat{x} - b_k + b_k - b_i) + \sum_{i=k+1}^{2k} (b_i - b_k + b_k - \hat{x}) =$$

$$= \sum_{i=1}^{k} (b_k - b_i) + k(\hat{x} - b_k) + \sum_{i=k+1}^{2k} (b_i - b_k) - k(\hat{x} - b_k) =$$

$$= \sum_{i=1}^{k} (b_k - b_i) + \sum_{i=k+1}^{2k} (b_i - b_k) .$$

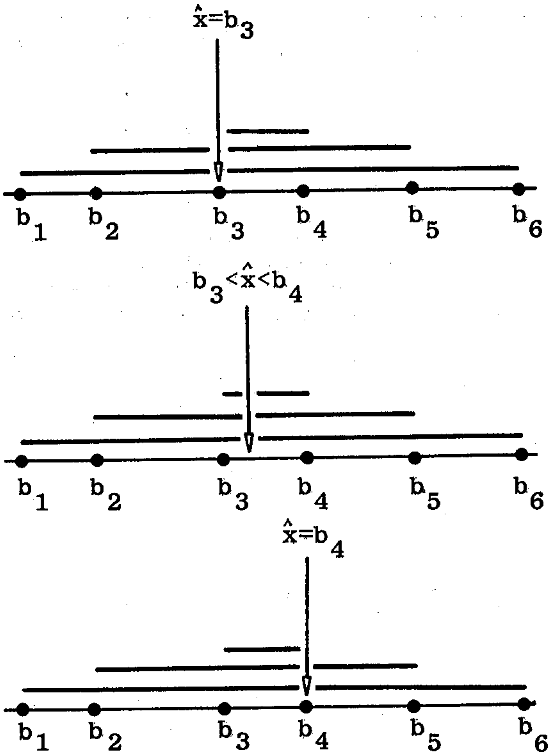


Σχήμα 2

Δηλαδή η ποσότητα $\varphi(\hat{x})$ είναι η ίδια για κάθε τιμή $\hat{x}$ στο διάστημα $b_k \le \hat{x} \le b_{k+1}$, και

$$\varphi(\hat{x}) = \varphi(b_k) = \varphi(b_{k+1})\,. \tag{27}$$

Οι παραπάνω σχέσεις γίνονται ευκολότερα κατανοητές με τη βοήθεια του σχήματος 3.



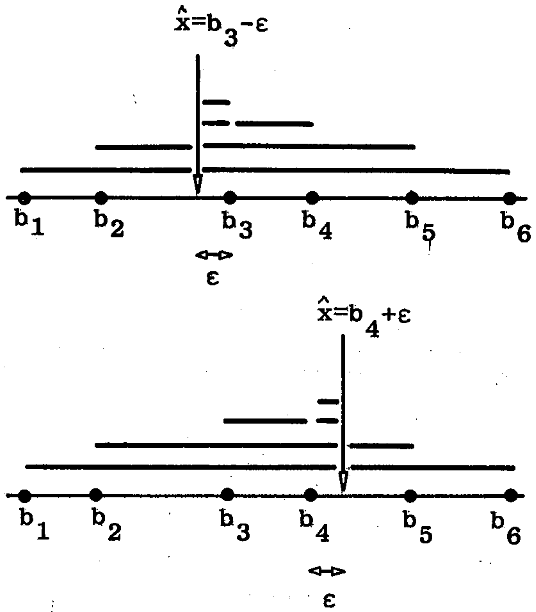
Σχήμα 3

Απομένει να δείξουμε ότι η ποσότητα φ αυξάνει όταν η τιμή $x=b_k$ ελαττωθεί, ή όταν η τιμή $x=b_{k+1}$ αυξηθεί, κατά μία μικρή θετική ποσότητα ε. Πράγματι

$$\varphi(b_k-\varepsilon) = \sum_{i=1}^{k-1} (b_k-\varepsilon-b_i) + [b_k-(b_k-\varepsilon)] + \sum_{i=k+1}^{2k} (b_i-b_k+\varepsilon) = \tag{28}$$

$$= \sum_{i=1}^{k-1} (b_k-b_i) - (k-1)\,\varepsilon + \varepsilon + \sum_{i=k+1}^{2k} (b_i-b_k) + k\varepsilon =$$

$$= \varphi(b_k) + 2\varepsilon\,,$$



Σχήμα 4

και

(29) $$\varphi(b_{k+1}+\varepsilon) = \sum_{i=1}^{k-1}(b_{k+1}+\varepsilon-b_i) + (b_{k+1}+\varepsilon-b_k) + (b_{k+1}+\varepsilon-b_{k+1}) + \sum_{i=k+2}^{2k}(b_i-b_{k+1}-\varepsilon) =$$

$$= \sum_{i=1}^{k-1}(b_{k+1}+\varepsilon-b_k+b_k-b_i) + 2\varepsilon + (b_{k+1}-b_k) + \sum_{i=k+2}^{2k}(b_i-b_k+b_k-b_{k+1}-\varepsilon) =$$

$$= \sum_{i=1}^{k-1}(b_k-b_i) + (k-1)(b_{k+1}+\varepsilon-b_k)+2\varepsilon + (b_{k+1}-b_k) + \sum_{i=k+2}^{2k}(b_i-b_k)-(k-1)(b_{k+1}+\varepsilon-b_k) =$$

$$= \sum_{i=1}^{k-1}(b_k-b_i)+2\varepsilon+\sum_{i=k+1}^{2k}(b_i-b_k) = \varphi(b_{k+1})+2\varepsilon .$$

Η παραπάνω απόδειξη γίνεται ευκολότερα κατανοητή με τη βοήθεια του σχήματος 4.

2.5. Σύγκριση των κριτήριων βέλτιστης εκτίμησης

Τα κριτήρια βέλτιστης εκτίμησης παραμέτρων που αναφέραμε δεν είναι τα μοναδικά, αλλά μόνο τα πιο συνηθισμένα. Μια οικογένεια κριτηρίων έχει τη μορφή

$$(30) \qquad \sum_{i=1}^{n} |v_i|^p = \min ,$$

όπου p θετικός ακέραιος, και περιλαμβάνει τα κριτήρια των ελαχίστων τετραγώνων (p=2) και της ελαχιστοποίησης του αθροίσματος των απόλυτων τιμών (p=1).

Με την εισαγωγή των κριτηρίων βελτιστοποίησης, το πρόβλημα της βέλτιστης εκτίμησης παραμέτρων δεν έχει λυθεί, αλλά έχει απλά αντικατασταθεί από ένα νέο πρόβλημα: Το πρόβλημα της επιλογής του «καλύτερου» κριτηρίου βελτιστοποίησης. Για να λυθεί το νέο αυτό πρόβλημα πρέπει πρώτα να ξεκαθαριστεί η έννοια του τί είναι καλύτερο, να δοθεί δηλαδή ένα μέτρο σύγκρισης ανάμεσα σε διάφορα προτεινόμενα κριτήρια. Κάτι τέτοιο μπορεί να κατορθωθεί μόνο μέσα από κάποια συγκεκριμένη θεωρία, σχετική με τη συμπεριφορά των σφαλμάτων v_i, αφού σ' αυτά αναφέρεται ουσιαστικά η βελτιστοποίηση.

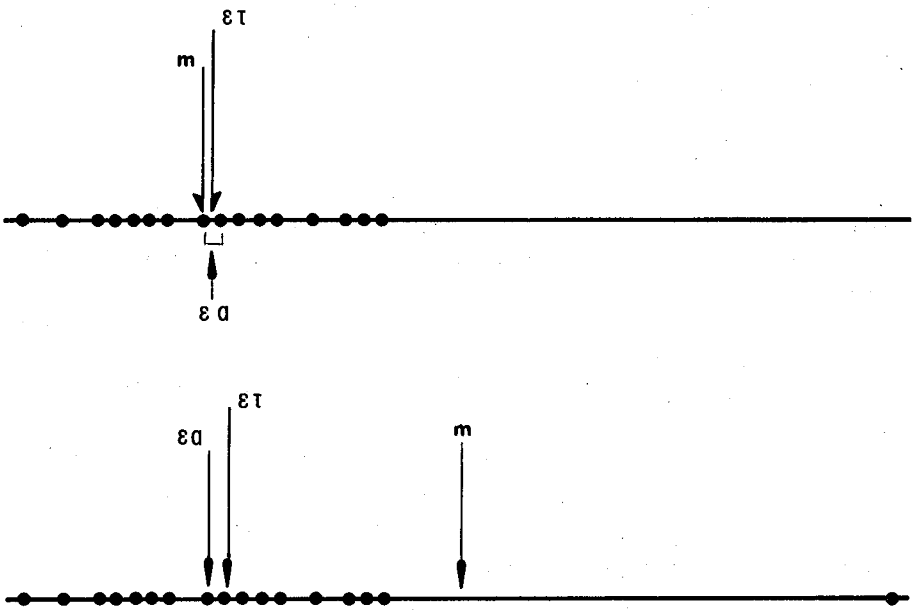
Μια τέτοια θεωρία θα εξετάσουμε στο επόμενο κεφάλαιο, όπου τα σφάλματα των παρατηρήσεων χαρακτηρίζονται σαν τυχαία σφάλματα και η συμπεριφορά τους περιγράφεται με τις μεθόδους της θεωρίας πιθανοτήτων και της μαθηματικής στατιστικής.

Εκείνο που μπορεί να γίνει εδώ είναι μόνο μια εμπειρική σύγκριση των κριτηρίων, σε σχέση με το συγκεκριμένο απλό παράδειγμα εκτίμησης παραμέτρων που εξετάζουμε, δηλαδή την εκτίμηση μιας παραμέτρου από πολλαπλές παρατηρήσεις. Απομένει να εξεταστούν, στα σχετικά κεφάλαια, οι τυχόν συνέπειες του περάσματος από την απλή αυτή περίπτωση σε περισσότερο πολύπλοκες περιπτώσεις εκτίμησης παραμέτρων.

Για μια εμπειρική σύγκριση των τριών κριτηρίων βελτιστοποίησης, θα εξετάσουμε τις τιμές που προκύπτουν για τη βέλτιστη εκτίμηση $\hat{x}$ μιας παραμέτρου μέσα από πολλαπλές παρατηρήσεις, σε τρία συγκεκριμένα παραδείγματα.

Παράδειγμα 1:

Στο πρώτο παράδειγμα (σχήμα 5), σε μια αρχική ομάδα παρατηρήσεων συγκεντρωμένων στην ίδια περιοχή τιμών, προστίθεται μια παρατήρηση που, για κάποιο λόγο (π.χ. χονδροειδές σφάλμα), διαφέρει σημαντικά από τις προηγούμενες. Ενώ η εκτίμηση $\hat{x}$ σύμφωνα με το



Σχήμα 5

κριτήριο των ελαχίστων τετραγώνων (ετ) και το κριτήριο των ελαχίστων απόλυτων τιμών (εα) αλλάζει πολύ λίγο, η επίδραση της νέας παρατήρησης πάνω στην εκτίμηση $\hat{x}$, σύμφωνα με το κριτήριο minmax (m), είναι εξαιρετικά μεγάλη. Στο παράδειγμα αυτό το κριτήριο minimax υστερεί σημαντικά σε σχέση με τα άλλα δύο κριτήρια.

Παράδειγμα 2:

Στο δεύτερο παράδειγμα (σχήμα 6) εξετάζονται δύο περιπτώσεις όπου το κριτήριο minimax οδηγεί στην ίδια εκτίμηση $\hat{x}$, ενώ οι παρατηρήσεις είναι σημαντικά μικρότερες στη δεύτερη περίπτωση σε σχέση με την πρώτη. Οι εκτιμήσεις σύμφωνα με τα άλλα δύο κριτήρια διαφέρουν σημαντικά στις δύο περιπτώσεις του παραδείγματος, αντανακλώντας τις