MINIMAXITÉ KAI ADMISSIBILITÉ

ΤΩΝ ΕΚΤΙΜΗΤΩΝ ΜΕ ΣΥΡΡΙΚΝΩΤΗ MATRICIEL

Περιεχόμενα

ΕΙΣΑΓΩΓΗ

Αυτή η εργασία αναφέρεται στον καθορισμό ορισμένων κριτηρίων επιλογής εκτιμητών με συρρικνωτή και αποτελεί μέρος μιας εκτεταμένης θεωρίας επί αυτών των εκτιμητών. Η θεωρία των εκτιμητών με συρρίκνωση, η οποία τοποθετείται στο πλαίσιο της Γραμμικής Στατιστικής Συμπερασματολογίας, εισάγεται από τους $James$ και $Stein$[1961], καθώς αυτοί επιβεβαιώνουν τη μη-παραδοχή του συνήθη εκτιμητή ελαχίστων τετραγώνων (ε.ε.τ) όταν η διάσταση του χώρου των παρατηρήσεων είναι ≥ 3, και έκτοτε έχει δώσει θέση σε πολυάριθμες και ποικίλες επεκτάσεις. Όμως, παρά τα πολλαπλά πλεονεκτήματα των εκτιμητών με συρρίκνωση, αυτοί δεν αναπληρώνουν πάντοτε τον ε.ε.τ.

Υιοθετώντας την $coordinate\ free$ προσέγγιση, η εργασία αυτή αποτελεί μια αναλυτική σύνθεση των άρθρων των $Fraisse$, $Raoult$, $Robert$ και Roy [$1987-1990$] στην οποία διευρύνεται το φάσμα εφαρμογής των εκτιμητών με συρρίκνωση (Κεφ.1) και δίνονται κριτήρια επιλογής μέσα σε κλάσεις, οι οποίες μερικές φορές κρίνονται πολύ γενικές (Κεφ.2,3,4).

1. ΤΟ ΜΟΝΤΕΛΟ

1.1

Θεωρούμε γενικώς μία τ.μ. y μέσα σ᾽ έναν διανυσματικό χώρο E, διάστασης n, επί του $\mathbb{R}$, η οποία ακολουθεί έναν κανονικό νόμο μέσου θ και διασποράς $\sigma^2 v$, όπου v μια γνωστή, συμμετρική, διγραμμική μορφή, θετικά ορισμένη επί του E^*, δυϊκός του E. Υποθέτουμε επιπλέον ότι το $\theta \in \Theta$, Θ διανυσματικός υπόχωρος του E διαστάσεως k $(3 \leq k \leq n-1)$ και ότι ο πολλαπλασιαστικός παράγοντας σ^2 είναι *άγνωστος*.

Σημειώνουμε με v_Θ^{-1} τον περιορισμό του v^{-1} επί του Θ, ϕ_0 τον ε.ε.τ του θ (v_Θ^{-1}-ορθογώνια προβολή του E επί του Θ) και $s^2(y)$ τον συνήθη εκτιμητή του σ^2.

Σ᾽ όλη την εργασία χρησιμοποιούμε την έννοια "$-$" για να σημειώσουμε την τετραγωνική μορφή που συνδέεται με μία συμμετρική, διγραμμική μορφή. Έτσι, για παράδειγμα,

$$s^2(y) = \frac{1}{n-k}\,\overline{v_{\Theta^\perp}^{-1}}(y - \phi_0(y))$$

($^\perp$ σημειώνει την ορθογωνικότητα αναφορικά με την v^{-1}).

Για μια συνάρτηση $g : E \to \mathbb{R}_+$ μετρήσιμη, σημειώνουμε:

$$\mathbb{E}_{\theta,\sigma}[g(y)] = \int_E g(y) P_{\theta,\sigma}(dy)$$

όπου $P_{\theta,\sigma} = N(\theta, \sigma^2 v)$, ο κανονικός νόμος μέσου θ και διασποράς $\sigma^2 v$.

Επειδή η v είναι θετικά ορισμένη, ο $N(\theta, \sigma^2 v)$ δέχεται πυκνότητα πιθανότητας, αναφορικά με το μέτρο του $Lebesgue$ λ_v επί του E, την συνάρτηση, από τον E στον $\mathbb{R}_+$,

$$y \rightsquigarrow (2\pi\sigma^2)^{-n/2} exp\Big\{-\frac{1}{2\sigma^2}\overline{v^{-1}}(y-\theta)\Big\}$$

Οι v^{-1}-ορθογώνιες προβολές του E επί του Θ και επί του $\Theta^{\perp}$ είναι ως γνωστό ανεξάρτητες τ.μ. (αναφορικά με τον $N(\theta, \sigma^2 v)$) και έχουμε:

$$N(\theta, \sigma^2 v) = N_{\Theta}(\theta, \sigma^2 v_{\Theta}) \times N_{\Theta^{\perp}}(0, \sigma^2 v_{\Theta^{\perp}})$$

Τούτο σημαίνει ότι:

$$\phi_0(y) \sim N_{\Theta}(\theta, \sigma^2 v_{\Theta}) \quad \text{και} \quad y - \phi_0(y) \sim N_{\Theta^{\perp}}(0, \sigma^2 v_{\Theta^{\perp}})$$

Οι εκτιμητές του θ συγκρίνονται μέσω της τετραγωνικής απώλειας συνδεδεμένη με μία θετική, διγραμμική, συμμετρική μορφή q επί του Θ. Όπως στους περισσότερους συγγραφείς, λαμβάνουμε και εδώ

$$\sigma^{-2}\overline{q}(\phi(y) - \theta)$$

σαν τιμή της απώλειας η οποία υφίσταται όταν εκτιμήσουμε το μέσο θ με το $\phi(y)$. Ο συνδεδεμένος με αυτήν κίνδυνος θα είναι λοιπόν

$$R_{\phi}(\theta, \sigma) = \sigma^{-2}\mathbb{E}_{\theta,\sigma}\Big[\overline{q}(\phi(y) - \theta)\Big]$$

Ένας εκτιμητής αναφοράς είναι ο ε.ε.τ, ϕ_0, του θ. Μέσω, λοιπόν, του παραπάνω κινδύνου συγκρίνουμε με τον ϕ_0 τους εκτιμητές ϕ τέτοιους ώστε το $\phi(y) - \phi_0(y)$ να είναι συγγραμμικό με την εικόνα του $\phi_0(y)$ υπό έναν ενδομορφισμό c του Θ. Είναι πρακτικό, αυτοί οι εκτιμητές να γραφούν υπό τη μορφή:

$$\phi(y) = \phi_0(y) - H\Big(\phi_0(y), y - \phi_0(y)\Big)c(\phi_0(y))$$

όπου $H : \Theta \times \Theta^{\perp} \to \mathbb{R}$ μετρήσιμη συνάρτηση καλούμενη *συρρικνωτής* ή *συνάρτηση συρρίκνωσης*.

Η ιδιαίτερη περίπτωση, όπου υπάρχει μία τετραγωνική μορφή $\overline{b}$ επί του Θ, τέτοια ώστε η $H(x, z)$ να είναι συνάρτηση του ζεύγους $\Big(\overline{b}(x), \overline{v_{\Theta^{\perp}}^{-1}}(z)\Big)$, αποτελεί το αντικείμενο αυτής της εργασίας, υπό διάφορα σχήματα υποθέσεων επί του c και επί του συρρικνωτή (και επομένως ιδιαίτερα επί του $\overline{b}$ και $\overline{v_{\Theta^{\perp}}^{-1}}$). Γενικά, οι υποθέσεις επί των αλγεβρικών τελεστών b και c είναι λιγότερο εξαναγκαστικές από τις υποθέσεις αναλυτικότητας για την συνάρτηση H, οι οποίες είναι περιοριστικές.

Ένα ιδιαίτερα περιοριστικό πλαίσιο επιλογής του b είναι εκείνο όπου διαθέτουμε λίγη πληροφορία για την διασπορά των παρατηρήσεων. Μια τέτοια περίπτωση είναι εκείνη όπου η κοινή διασπορά των παρατηρήσεων είναι πλήρως άγνωστη. Η συνάρτηση συρρίκνωσης εξαρτάται τότε μόνο από το $\phi_0(y)$, μέσω της τιμής στο $\phi_0(y)$ της αντίστροφης τετραγωνικής μορφής του εκτιμητή της διασποράς η οποία ακολουθεί την κατανομή $Wishart$.

Σε μερικές περιπτώσεις, η επιλογή των b και c είναι πιο περιοριστική από αυτή του γενικού μοντέλου, όπως όταν $b = v_{\Theta}^{-1}$ και $c = id_{\Theta}$ για πολυάριθμα αποτελέσματα (βλ. επίσης Κεφ.4) ή όταν $c = v_{\Theta}b$ στο Κεφ.2.

1.2. Εκτιμητής των James – Stein

Είναι καλά γνωστό ότι ο ε.ε.τ, ϕ_0, είναι ελαχίστου κινδύνου μεταξύ των αμερολήπτων γραμμικών εκτιμητών (Θεώρημα των $Gauss - Markov$) και ακόμη, για τον κανονικό νόμο, μεταξύ όλων των αμερολήπτων εκτιμητών (Θεώρημα $Lehmann - Scheff\acute{e}$). Επίσης, αυτός είναι *minimax* (δηλ. ο ϕ_0 ελαχιστοποιεί το μέγιστο κίνδυνο, $\sup_{\theta,\sigma} R_\phi(\theta,\sigma)$, μεταξύ όλων των εκτιμητών ϕ) και *σταθερού κινδύνου* ίσου με $tr(v_\Theta q)$. Για τον κανονικό νόμο, ο ϕ_0 είναι εξίσου ο εκτιμητής *μεγίστης πιθανοφάνειας*.

Όμως, το 1956, ο *Stein* αποδεικνύει ότι ο ϕ_0 είναι *παραδεκτός* μόνο αν $k \leq 2$, δηλαδή, όταν $k \geq 3$, υπάρχει πάντοτε ένας εκτιμητής $\widetilde{\phi}$ τέτοιος ώστε:

$$R_{\widetilde{\phi}}(\theta,\sigma) \leq R_\phi(\theta,\sigma)$$

για κάθε (θ,σ), με την αυστηρή ανισότητα να ισχύει για ένα τουλάχιστο (θ,σ).

Η απόδειξη του, για το μέρος της "παραδοχής", όταν $k \leq 2$, είναι όμοια με εκείνη του Κεφ.3, η οποία αποδεικνύει μία αναγκαία και ικανή συνθήκη ύπαρξης εκτιμητών με συρρικνωτή κυριαρχώντας ομοιομόρφως τον ε.ε.τ για μια κλάση απωλειών.

Αυτή η απόδειξη δεν ήταν εποικοδομιτική, όμως αργότερα οι *James* και *Stein* [1961] επιδεικνύουν έναν εκτιμητή κινδύνου ομοιομόρφως μικρότερου από εκείνον του ϕ_0, για τον "συνήθη" κίνδυνο συνδεδεμένο με την $q = v_\Theta^{-1}$. Αυτός γράφεται υπό τη μορφή:

$$\phi_{JS}(y) = \left(1 - \rho\, \frac{s^2(y)}{\overline{\overline{v_\Theta^{-1}(\phi_0(y))}}}\right)\phi_0(y)$$

όπου $\rho = \frac{(k-2)(n-k)}{n-k+2}$. Φαίνεται τότε το γιατί ο εκτιμητής ϕ_{JS}, λεγόμενος των $James - Stein$, ονομάζεται *εκτιμητής με συρρικνωτή*(ή εκτιμητής με συρρικνωτή κλίμακας(*scalaire*)) : αυτός πολλαπλασιάζει τον ε.ε.τ με ένα συντελεστή μικρότερο του 1 και τον συρρικνώνει προς το 0.

2. ΟΙ ΕΚΤΙΜΗΤΕΣ ΜΕ ΣΥΡΡΙΚΝΩΤΗ

2.1

Αυτοί αποτελούν στη συνέχεια το αντικείμενο μελέτης πολλών συγγραφέων, οι οποίοι εκμεταλεύονται το αποτέλεσμα των *James* και *Stein* και το γενικεύουν κατά διάφορους τρόπους για να σχηματίσουν σιγά-σιγά έναν ειδικό κλάδο της γραμμικής στατιστικής (βλ. π.χ.*Judge* και *Bock*[1987]). Σε γενικές γραμμές, η πορεία της εξέλιξης των εκτιμητών με συρρικνωτή έχει ως εξής:

Οι *Baranchik*[1964] και *Bock*[1975] γενικεύουν τα αποτελέσματα των *James* και *Stein*[1961], ως προς την έννοια της συνάρτησης συρρικνώσεως, θεωρώντας εκτιμητές της μορφής:

$$\phi_B(y) = \left(1 - h\left(\frac{s^2(y)}{\overline{\overline{v_\Theta^{-1}(\phi_0(y))}}}\right)\right)\phi_0(y)$$

και αποδεικνύοντας επί της συναρτήσεως h επαρκείς συνθήκες ομοιόμορφης κυριαρχίας του ϕ_B επί του ϕ_0. Διάφοροι συγγραφείς (όπως π.χ. οι *Strawderman*[1973], *Efron* και *Morris*[1976]) γενικεύουν προοδευτικά αυτές τις υποθέσεις.

Μια δεύτερη γενίκευση ακολουθεί από τους *Berger* και *Bock*[1976] με την αντικατάσταση του συντελεστή συρρίκνωσης από τον *ενδομορφισμό συρρίκνωσης* c, ενδομορφισμός επί του Θ, εναρμονίζοντας την συρρίκνωση κατά συνιστώσα. Αυτοί θεωρούν τους εκτιμητές:

$$\phi_{BB}(y) = \left(id_\Theta - h\left(\frac{s^2(y)}{\overline{\overline{v_\Theta^{-1}(\phi_0(y))}}}\right)c\right)\phi_0(y)$$

όπου id_Θ ο ταυτοτικός ενδομορφισμός επί του Θ.

Οι $Judge$ και $Bock$[1978] γενικεύουν αυτούς τους εκτιμητές στους:

$$\phi_{JB}(y) = \left(id_\Theta - h\Big(\frac{s^2(y)}{\overline{b}(\phi_0(y))}\Big)c\right)\phi_0(y)$$

όπου b διγραμμική, συμμετρική μορφή, επί του Θ, η οποία διαγωνιοποιείται, καθώς και ο ενδομορφισμός c, μέσα στην ίδια βάση η οποία είναι v_Θ^{-1}-ορθοκανονική και q-ορθογώνια. Επειδή οι συναρτήσεις συρρικνώσεως έχουν θετικές τιμές, παρατηρούμε ότι, μέσα σε μία τέτοια βάση, οι διάφορες συντεταγμένες του $\phi_0(y)$ "συρρικνώνονται" προς το 0. Αυτοί οι εκτιμητές θα ονομασθούν εξίσου *εκτιμητές με συρρικνωτή* και συγκεκριμένα $matriciels$, υποδεικνύοντας ότι η συρρίκνωση μεταβάλλεται ακολουθώντας τις συντεταγμένες.

Οι επαρκείς συνθήκες επί της h, θεωρούμενες από τους $Judge$ και $Bock$, ώστε ο ϕ_{JB} να κυριαρχεί τον ϕ_0, γενικεύονται από τους $Cellier$ και $Fourdrinier$[1985] στην περίπτωση μη-διαφορίσιμων συναρτήσεων συρρίκνωσης. Αυτοί επεκτείνουν τα αποτελέσματα τους θεωρώντας τους εκτιμητές της μορφής:

$$\phi(y) = \left(id_\Theta - h\big(\overline{b}(\phi_0(y)), s^2(y)\big)c\right)\phi_0(y)$$

με υποθέσεις πολύ λιγότερο περιοριστικές επί της συνάρτησης συρρίκνωσης και με παρατήρηση y ακολουθώντας ένα νόμο με ελλειπτική συμμετρία (η περίπτωση μιας κανονικής παρατήρησης δίδεται από τους $Cellier$, $Fourdrinier$ και $Robert$[1989]).

Επεκτείνοντας τα αποτελέσματα του $Berger$[1976], οι $Fraisse$, $Robert$ και Roy[1987] (Κεφ.1) θεωρούν τους εκτιμητές ϕ της ανωτέρω μορφής συνδεδεμένους με διαφορίσιμες συναρτήσεις συρρικνώσεως και δίνουν επαρκείς συνθήκες κυριαρχίας επί του ϕ_0, στην περίπτωση που οι διάφοροι διανυσματικοί τελεστές (ο ενδομορφισμός συρρίκνωσης c, η διγραμμική μορφή b που υπεισέρχεται στην συνάρτηση συρρικνώσεως) δεν διαγωνιοποιούνται στην ίδια βάση του Θ, όπως οι q και v_Θ^{-1}. Οι τοιουτοτρόπως λαμβανόμενες συνθήκες χρησιμοποιούνται στο Κεφ.4 για την σύγκριση των εκτιμητών με συρρίκνωση.

Βεβαίως, υπάρχουν γενικεύσεις των ανωτέρω αποτελεσμάτων για μη-κανονικούς νόμους. Ο $Brown$[1966] γενικεύει το αποτέλεσμα του $Stein$[1956] με την εκτίμηση της παραμέτρου θέσεως μιας πολύ γενικής κλάσεως νόμων. Οι $Brandwein$ και $Strawderman$[1991] γενικεύουν τα αποτελέσματα του $Stein$[1981], ο οποίος εισάγει μια πολύ γενική κλάση εκτιμητών με συρρικνωτή με επαρκείς συνθήκες κυριαρχίας επί του ϕ_0 πολύ εξαναγκαστικές, σ' ένα πλαίσιο σφαιρικά συμμετρικών νόμων δεδομένης πυκνότητας, ενώ οι $Cellier$ και $Fourdrinier$[1994] γενικεύουν τα αποτελέσματα εκείνων σ' ένα γενικότερο πλαίσιο νόμων με ελλειπτική συμμετρία (χωρίς την ανάγκη ύπαρξης πυκνότητας) και χωρίς τις ισχυρές αναλυτικές υποθέσεις εκείνων.

Να σημειώσουμε ότι για τους περισσότερους νόμους, ο θεωρούμενος εκτιμητής ϕ_0 δεν είναι εκείνος των ελαχίστων τετραγώνων αλλά εκείνος της μεγίστης πιθανοφάνειας. Είναι λοιπόν ενδιαφέρον να εξετάσουμε την επιρροή των εκτιμητών με συρρικνωτή επί της επίδοσης αυτού του εκτιμητή (Κεφ.4).

Στην συνέχεια του $Berger$[1978], ο οποίος μελετά τις πολυωνυμικές απώλειες, οι $Brandwein$ και $Strawderman$[1980] αποδεικνύουν εξίσου αποτελέσματα κυριαρχίας για κυρτές απώλειες. Ένα κριτήριο σύγκρισης, ενδιαφέρον για πολλούς συγγραφείς, είναι εκείνο του "μέσου τετραγωνικού σφάλματος" το οποίο θεωρεί τον πίνακα της δεύτερης ροπής του εκτιμητή από τον οποίον αφαιρείται το θ. Στην πράξη, αν ένας εκτιμητής ϕ κυριαρχεί έναν εκτιμητή ϕ' με την έννοια αυτού του κριτηρίου (δηλ. ότι η διαφορά των δύο προαναφερθέντων πινάκων είναι θετική) ο ϕ θα κυριαρχεί τον ϕ' για όλους τους κινδύνους συνδεδεμένους με τετραγωνικές μορφές.

Στην περίπτωση που η διασπορά είναι πλήρως γνωστή, ο $Brown$[1975] αποδεικνύει ότι δεν υπάρχει εκτιμητής με συρρικνωτή καλύτερος του ϕ_0 με την έννοια αυτού του κριτηρίου. Ο $C.Robert$[1987]([41]) γενικεύει αυτό το αποτέλεσμα στην περίπτωση που η διασπορά είναι γνωστή εκτός από έναν πολλαπλασιαστικό παράγοντα (Κεφ.3).

2.2. Βελτίωση και παραδοχή

Όπως είδαμε στην §2.1, οι κλάσεις των εκτιμητών με συρρικνωτή αγγίζουν μία πολύ μεγάλη γενικότητα, η οποία στην πράξη φαίνεται να είναι ένα εμπόδιο στην χρήση αυτών των εκτιμητών, καθώς ο χρήστης δεν διαθέτει κριτήρια επιλογής μέσα στην προτεινόμενη κλάση εκτός από τις επαρκείς συνθήκες κυριαρχίας.

Επεκτείνοντας τις χρησιμοποιούμενες μεθόδους, ο $C.Robert$[1987]([41]), για να εγγυηθεί μια ομοιόμορφη κυριαρχία επί του ϕ_0, εξασφαλίζει βελτιώσεις των υπαρχόντων εκτιμητών και κριτηρίων που είναι πιο ακριβείς από την επιλογή των εκτιμητών με συρρικνωτή. Στο Κεφ.4, λοιπόν, μελετάμε τις συνθήκες υπό τις οποίες μπορούμε να βελτιώσουμε έναν δεδομένο εκτιμητή με συρρικνωτή από έναν εκτιμητή "κλίμακας", τροποποιώντας τον ενδομορφισμό ή τη συνάρτηση συρρικνώσεως. Ακολουθώντας αυτές τις τεχνικές, οι $Judge$ και $Bock$[1978] επεκτείνουν τον "*positive rule*", εισαγόμενο από τον $Baranchik$[1964], στο σύνολο των εκτιμητών τους. Συγκεκριμένα, αυτοί δείχνουν ότι, στην πράξη είναι προτιμότερο, όταν θεωρείται ο εκτιμητής κατά συντεταγμένη, να πραγματοποιείται μια αληθινή συρρίκνωση: "κολόβωση του συντελεστή συρρίκνωσης στο 0". Ο $Stein$[1981] και οι Dey και $Berger$[1983] βελτιώνουν υπό κολόβωση τον εκτιμητή των $James - Stein$. Μια τέτοια κολόβωση παρουσιάζεται στο τέλος του Κεφ.4.

Επειδή ο εκτιμητής ϕ_0 είναι μη-παραδεκτός, παρουσιάζει ενδιαφέρον να αναρωτηθούμε για την παραδοχή των θεωρούμενων εκτιμητών γενικά. Γενικεύοντας το αποτέλεσμα του $Stein$[1955], ο $Farrell$[1968] αποδεικνύει μία ικανή και αναγκαία συνθήκη για την παραδοχή σ' ένα πλαίσιο εξαιρετικά γενικό. Ακολουθώντας τον $Brown$[1971], οι $Berger$ και $Srinivasan$[1978] αποδεικνύουν, χάρις σ' αυτή τη συνθήκη, ότι οι γενικευμένοι εκτιμητές του $Bayes$ των φυσικών παραμέτρων μιας εκθετικής οικογένειας αποτελούν μια πλήρη κλάση. Τα αποτελέσματα αυτά επαναλαμβάνονται στον $Brown$[1986]. Μια ειδική περίπτωση αυτών είναι εκείνη της εκτίμησης του μέσου μιας πολυδιάστατης κανονικής κατανομής, όπου η διασπορά είναι πλήρως γνωστή. Για την περίπτωση αυτή, συμπεραίνεται μία αναγκαία συνθήκη της παραδοχής των εκτιμητών με συρρικνωτή αυτού του μέσου. Οι $Fraisse$, $Raoult$, $Robert$ και Roy[1990] γενικεύουν τα παραπάνω αποτελέσματα αποδεικνύοντας ότι για τις ενδιαφέρουσες παραμέτρους, οι οποίες δεν είναι οι φυσικοί παράμετροι, οι παραδεκτοί εκτιμητές είναι πάντοτε όρια εκτιμητών του $Bayes$ και επιδεικνύεται μία αναγκαία μορφή αυτών των εκτιμητών. Μία ειδική περίπτωση της κατάστασης αυτής είναι η εκτίμηση του μέσου ενός πολυδιάστατου κανονικού νόμου, όπου η διασπορά είναι άγνωστη κατά έναν πολλαπλασιαστικό παράγοντα που παίρνει τιμές σ' ένα συμπαγές διάστημα του $(0, +\infty)$. Σ' αυτή την περίπτωση δίνουν μία αναγκαία συνθήκη παραδοχής των εκτιμητών με συρρίκνωση, κατά την οποία μέσα σε μία βάση v_Θ^{-1}-ορθοκανονική οι πίνακες των b και c οφείλουν να είναι ανάλογοι. Αυτή η γενίκευση αποτελεί το αντικείμενο του Κεφ.2.

Ο δρόμος προς την παραδοχή, όπως φαίνεται στο Κεφ.2, περνά από μια *bayesienne* μεταχείρηση. Πράγματι, αυτή η μεταχείρηση επιτρέπει να εκφράσουμε ταυτόχρονα το "φαινόμενο *Stein*" (δηλ. το γιατί η ταυτόχρονη εκτίμηση ανεξαρτήτων προβλημάτων βελτιώνει συνολικά αυτή την εκτίμηση(π.χ $Berger$[1980])) και να προτείνουμε εκτιμητές ταυτοχρόνως καλύτερους από τον ε.ε.τ (π.χ $Berger$[1982], $Berger$ και $Berliner$[1984]) και μάλιστα παραδεκτούς (π.χ $Alam$[1973], $Berger$[1980])

Στο Κεφ.2 δεν αντιμετωπίζουμε τους εκτιμητές με συρρίκνωση υπό την *bayesienne* οπτική, εκεί οι εκτιμητές του *Bayes* φαίνονται ως εργαλείο. Για μια *bayesienne* προσέγγιση των εκτιμητών με συρρίκνωση του Κεφ.1, βλέπε *Critikou* και *Terzakis*[1992].

Σχεδιάγραμμα της εργασίας

Σκοπός αυτής της εργασίας, πέρα από το να καταστήσει σαφείς κάποιες υποθέσεις κυριαρχίας του ϕ_0 από τον ϕ φερόμενες επί της συνάρτησης συρρικνώσεως h (εξαρτόμενη από τα $\overline{b}(x)$ και $\overline{v^{-1}}(z)$) και των αλγεβρικών τελεστών q, b και c, είναι να δώσει και κριτήρια επιλογής στο εσωτερικό της κλάσης των εκτιμητών ϕ.

Συγκεκριμένα:

– Στο Κεφ.1 θεωρείται το πρόβλημα της εκτίμησης του μέσου ενός πολυδιάστατου κανονικού νόμου, όταν η διασπορά είναι άγνωστη κατά έναν πολλαπλασιαστικό παράγοντα, από εκτιμητές με συνάρτηση συρρικνώσεως *διαφορίσιμη*, η οποία μας επιτρέπει τη χρήση των μεθόδων ολοκλήρωσης κατά μέρη εισαγόμενη από τον *Stein*[1973, 1981]. Ο υπολογισμός του κινδύνου των θεωρούμενων εκτιμητών οδηγεί σε αμερόληπτους εκτιμητές αυτού του κινδύνου, από τους οποίους συμπεραίνονται ικανές συνθήκες κυριαρχίας του ε.ε.τ οι οποίες καθίστανται πιο σαφείς αν περιορίσουμε την κλάση των ενδομορφισμών συρρίκνωσης.

– Στο Κεφ.2 μελετάται αρχικά η εκτίμηση των μη-φυσικών παραμέτρων σε εκθετικές οικογένειες πυκνοτήτων και αποδεικνύεται, υπό μία υπόθεση συμπαγότητας, ότι οι παραδεκτοί εκτιμητές αυτών των παραμέτρων είναι όρια εκτιμητών του *Bayes* και μπορούν να εκφρασθούν υπό μια μορφή ιδιαίτερα συναρτησιακή. Στο πλαίσιο αυτής της μελέτης, μια ενδιαφέρουσα ειδική περίπτωση οδηγεί στο πρόβλημα εκτίμησης του Κεφ.1, συμπεραίνοντας από το γενικό αποτέλεσμα μία αναγκαία συνθήκη για την παραδοχή των εκτιμητών με συρρικνωτή.

– Στο Κεφ.3 δίδεται μια αναγκαία και ικανή συνθήκη κυριαρχίας του ε.ε.τ του κανονικού μέσου γενικεύοντας το αποτέλεσμα του *Brown*[1975] και εν μέρει το αποτέλεσμα του *C.Robert*[1987] ([41]) αντιστοίχως στις περιπτώσεις που $b = v_\Theta^{-1}$ και b οποιαδήποτε διγραμμική μορφή.

– Στο Κεφ.4 επιτυγχάνεται "βελτίωση" των εκτιμητών των προηγούμενων κεφαλαίων ενισχύοντας αρχικά την συρρίκνωση τους μέσω του ενδομορφισμού συρρικνώσεως θεωρώντας δεδομένη τη συνάρτηση συρρίκνωσης και στην συνέχεια τροποποιώντας ή κόβοντας την συνάρτηση συρρικνώσεως τους. Ο όρος "βελτίωση" σημαίνει την εξασφάλιση, κάτω από ορισμένες συνθήκες, μιας ομοιόμορφης ως προς (θ, σ) ελάττωσης του κινδύνου του συνδεδεμένου με αυτόν εκτιμητή. Μια μέθοδος στο τέλος του κεφαλαίου αυτού μας επιτρέπει να εκτιμήσουμε την τιμή ενός εκτιμητή ϕ^1 (με την έννοια ενός δεδομένου τετραγωνικού κινδύνου) εκτιμώντας, με προσομοίωση, το άνω φράγμα του κέρδους το οποίο μπορεί να επιφέρει ένας εκτιμητής με συρρικνωτή αναφορικά με τον ε.ε.τ. Μπορούμε, λοιπόν, να συγκρίνουμε (πάντοτε με προσομοίωση) το κέρδος του εκτιμητή ϕ^1 με το φράγμα αυτό.

– Η υπόθεση της διαφορισιμότητας στην συνάρτηση συρρίκνωσης μας επιτρέπει τη χρήση του Λήμματος του *Stein*, η δε χρήση των αλγεβρικών τελεστών δημιουργεί την ανάγκη ενός αλγεβρικού λήμματος. Και τα δύο δίδονται ως Παράρτημα στο Κεφ.1. Όμως, στο τέλος της εργασίας, ακολουθεί ένα άλλο Παράρτημα, το οποίο δίδεται για να διευκρινισθούν κλασσικά αποτελέσματα που εμφανίζονται στους υπολογισμούς και αναφέρεται στην τεχνική της "*poissonisation*".

Κεφάλαιο 1

ΕΚΤΙΜΗΤΕΣ ΜΕ ΔΙΑΦΟΡΙΣΙΜΟ ΣΥΡΡΙΚΝΩΤΗ ΓΙΑ ΜΙΑ ΤΕΤΡΑΓΩΝΙΚΗ ΣΥΝΑΡΤΗΣΗ ΑΠΩΛΕΙΑΣ

1.1

Στο Κεφάλαιο αυτό, ακολουθώντας τους *Fraisse*, *Robert* και *Roy*[1987], θεωρούμε το πρόβλημα της εκτίμησης του μέσου ενός πολυδιάστατου κανονικού νόμου, όταν η διασπορά είναι γνωστή εκτός από έναν πολλαπλασιαστικό παράγοντα, υπό μία γενική τετραγωνική συνάρτηση απώλειας, δια μέσου των εκτιμητών με συρρικνωτή έχοντας διαφορίσιμη συνάρτηση συρρίκνωσης κάτω από ασθενείς αλγεβρικές υποθέσεις. Οι εν λόγω εκτιμητές, υπό ικανές συνθήκες, κυριαρχούν ομοιόμορφα τον ε.ε.τ. Οι εκτιμητές με συρρίκνωση που θεωρούμε ανήκουν σε μια αισθητά μεγαλύτερη κλάση εκτιμητών από αυτές που υπάρχουν στο μεγαλύτερο μέρος της βιβλιογραφίας και αφορούν το ίδιο πρόβλημα.

1.2 ΤΟ ΜΟΝΤΕΛΟ

1.2.1 Έννοιες

Έστω V ένας διανυσματικός χώρος, διάστασης n επί του $\mathbb{R}$ και a μια διγραμμική μορφή επί του V. Σημειώνουμε με $\bar{a}$ την τετραγωνική μορφή που συνδέεται με την a.

Άν a είναι μια διγραμμική συμμετρική μορφή επί του V, μπορούμε να την θεωρήσουμε σαν έναν ομομορφισμό από τον V στον V^* (όπου V^* ο δυϊκός του V). Σημειώνουμε a^- έναν γενικευμένο αντίστροφο ($g - inverse$) του a, δηλ. έναν ομομορφισμό του V^* με τον V που ικανοποιεί τη σχέση : $a\, a^- a = a$. Έχουμε λοιπόν, $Im(a^-) \oplus Ker\, a = V$.

1.2.2

Έστω y μια τυχαία μεταβλητή με τιμές σε έναν διανυσματικό χώρο E, επί του $\mathbb{R}$, διάστασης n, η οποία ακολουθεί έναν κανονικό νόμο $N(\theta, \sigma^2 v)$, όπου το θ ανήκει σε έναν διανυσματικό υπόχωρο Θ του E, διάστασης k (με $k < n$), v είναι μια γνωστή, θετικά ορισμένη, συμμετρική, διγραμμική μορφή επί του E^* και οι παράμετροι θ και σ ($\sigma > 0$) είναι άγνωστες.

Οι θεωρούμενοι εκτιμητές του θ, δηλ. οι απεικονίσεις ϕ από τον E στον Θ, συγκρίνονται, μέσω μιας τετραγωνικής απώλειας, ορισμένη από μια θετική, συμμετρική, διγραμμική μορφή q επί του Θ. Θεωρούμε αυτή τη συνάρτηση απώλειας να είναι η: $\sigma^{-2}\bar{q}(\phi(y) - \theta)$. Ο κίνδυνος του

ϕ είναι λοιπόν:

$$R_\phi(\theta,\sigma) = \sigma^{-2}\mathbb{E}_{\theta,\sigma}[\bar{q}(\phi(y)-\theta)]$$

όπου για κάθε μετρήσιμη απεικόνιση $f : E \to \mathbb{R}$, σημειώνουμε:

$$\mathbb{E}_{\theta,\sigma}[f(y)] = \int_E f(y)N(\theta,\sigma^2 v; dy)$$

Σημειώνουμε με ϕ_0 τον εκτιμητή μεγίστης πιθανοφάνειας του θ, δηλ. την v^{-1}-ορθογώνια προβολή επί του Θ. Ο νόμος του $\phi_0(y)$ είναι $N(\theta,\sigma^2 v_\Theta)$, όπου v_Θ^{-1} είναι ο περιορισμός του v^{-1} επί του Θ. Αυτός ο εκτιμητής είναι *minimax* και δέχεται σταθερό κίνδυνο ίσο με : $R_{\phi_0}(\theta,\sigma) = tr(v_\Theta\, q)$.

Παρατήρηση 1.2.1
Το μοντέλο αυτό είναι στην πραγματικότητα το μοντέλο της γραμμικής παλινδρόμησης. Πράγματι, αν $x_1, x_2, \dots, x_k$ είναι k διανύσματα του E, που παράγουν τον Θ, για $\theta = \sum_{i=1}^k \beta_i x_i$ έχουμε $y = \sum_{i=1}^k \beta_i x_i + u$, όπου u ακολουθεί έναν κανονικό νόμο, μέσου 0 και διασποράς $\sigma^2 v$. Η εκτίμηση του $(\beta_1, \beta_2, \dots, \beta_k)$ μας οδηγεί στην εκτίμηση του θ μέσω της βάσης $(x_1, x_2, \dots, x_k)$ του Θ. Όμως, χωρίς να επιλέξουμε μια βάση αναφοράς για τον Θ καταλήγουμε σε πιο συνοπτικές εκφράσεις των αποτελεσμάτων.

1.2.3 Οι εκτιμητές με συρρικνωτή

Έστω b μια θετική, συμμετρική, διγραμμική μορφή επί του Θ, c ένας ενδομορφισμός του Θ και h μια απεικόνιση από τον $(\mathbb{R}_+)^2$ στον $\mathbb{R}_+$. Ο c ονομάζεται *ενδομορφισμός συρρίκνωσης* και η h *συνάρτηση συρρίκνωσης*. Θεωρούμε την υπόθεση H_1,

ΥΠΟΘΕΣΗ $\mathbf{H_1}$: $Ker\ b \subset Ker\ c$

Οι *εκτιμητές με συρρικνωτή* θα είναι της μορφής :

$$\phi(y) = \Big(id_\Theta - h\Big(\bar{b}(\phi_0(y)), (n-k)/\overline{v^{-1}}(y-\phi_0(y))\Big)c\Big)\phi_0(y) \tag{1.1}$$

Σε όλα τα παρακάτω θα σημειώνουμε :

$$t = \bar{b}(\phi_0(y)) \quad \text{και} \quad s^2 = \frac{\overline{v^{-1}}(y-\phi_0(y))}{n-k}$$

s^2 ο συνήθης εκτιμητης της διασποράς σ^2.
Θεωρούμε μια βάση του E, $(e_1, e_2, \dots, e_n)$, v^{-1}- ορθοκανονική, τέτοια ώστε τα $(e_1, e_2, \dots, e_k)$ να σχηματίζουν μια b-ορθογώνια βάση για τον Θ, τέτοια ώστε αν $m = rg(b)$, τα $k-m$ τελευταία διανύσματα αυτής της βάσης να παράγουν τον $Ker\,b$. Έχουμε τότε:

$$\phi_0(y) = \sum_{i=1}^k y_i e_i \quad , \quad \theta = \sum_{i=1}^k \theta_i e_i$$

$$t = \sum_{i=1}^m b_i y_i^2 \quad \text{και} \quad s^2 = \frac{1}{n-k} \sum_{i=k+1}^n y_i^2$$

απ′ όπου προκύπτει :

$$\phi(y) = \sum_{i=1}^{k} y_i e_i \ - \ h\Big(\sum_{i=1}^{m} b_i y_i^2 \ , \ (n-k)(\sum_{i=k+1}^{n} y_i^2)^{-1} \Big) \sum_{i=1}^{k} \sum_{j=1}^{k} c_{ij} y_j e_i$$

όπου c_{ij} τα στοιχεία του πίνακα που συνδέεται με τον ενδομορφισμό c.

Οι εκτιμητές της μορφής (1.1) γενικεύουν αυτούς των $Berger$ και $Bock$[1976] , $Cellier$ και $Fourdrinier$[1985] , $Efron$ και $Morris$[1976] , $Judge$ και $Bock$[1978] , $Strawderman$[1973] οι οποίοι θεωρούν μόνο συναρτήσεις συρρίκνωσης μίας μεταβλητής ή υποθέτουν ότι τα b, c και q είναι ταυτοχρόνως διαγωνιοποιήσιμα μέσα σε μια βάση v_Θ^{-1}-ορθοκανονική . Οι $Cellier$ και $Fourdrinier$[1985] εισάγουν τους εκτιμητές με συνάρτηση συρρίκνωσης όχι αναγκαία διαφορίσιμη (ούτε καν συνεχή) ενώ εδώ υποθέτουμε, ότι η h δέχεται μερικές παραγώγους αναφορικά και με τις δύο μεταβλητές της. Χάρη σε αυτήν την αναλυτική υπόθεση, καταργούνται, αναφορικά με τα παραπάνω άρθρα,οι αρκετά ισχυροί περιορισμοί επί της φασματικής ανάλυσης των b, c και q.

Θα λέμε ότι ο ϕ είναι *ομαλός* ($regulier$) αν η συνάρτηση συρρίκνωσής του, h, ικανοποιεί τις παρακάτω συνθήκες (αν με h_i' συμβολίσουμε τη μερική παράγωγο της h ως προς την i-μεταβλητή):

$$\mathbb{E}_{\theta,\sigma}[\, t\, h^2(t, 1/s^2)\,] \quad , \quad \mathbb{E}_{\theta,\sigma}\, [\, t\, h_1'(t, 1/s^2)\,] \quad ,$$

$$\mathbb{E}_{\theta,\sigma}[\, \frac{t}{s^4}\, h(t, 1/s^2)\, h_2'(t, 1/s^2)\,] \quad \text{και} \quad \mathbb{E}_{\theta,\sigma}[\, \frac{t}{s^2}\, h^2(t, 1/s^2)\,]$$

είναι πεπερασμένες, για κάθε $\theta \in \Theta$ και $\sigma > 0$.

Θα δούμε ότι, αυτοί οι περιορισμοί εξασφαλίζουν το πεπερασμένο του κινδύνου του εκτιμητή που συνδέεται με την h (βλ. Πρόταση 1.3.1) και την εγκυρότητα των ολοκληρώσεων κατά μέρη που γίνονται σε διάφορες αποδείξεις για την εφαρμογή του Λήμματος του $Stein$ (βλ.Λήμμα 1.6.2).

Η b δεν είναι απαραίτητα θετικά ορισμένη, αλλά υποθέτουμε οτι $rg(b) \geq 3$. Αυτός ο περιορισμός είναι φυσικός αφού, όταν η H_1 ικανοποιείται και αν $rg(b) \leq 2$, είναι φανερό οτι ο εκτιμητής του τύπου (1.1) διαφέρει από τον ϕ_0 μόνο κατά δύο το πολύ ανεξάρτητες κατευθύνσεις. Ο $Stein$[1956] έδειξε οτι, για αυτήν την περίπτωση, ο ϕ_0 δεν μπορεί να βελτιωθεί ομοιομόρφως από τον ϕ.

1.3 ΚΙΝΔΥΝΟΣ ΤΩΝ ΕΚΤΙΜΗΤΩΝ ΜΕ ΣΥΡΡΙΚΝΩΤΗ

1.3.1

Σύμφωνα με την υπόθεση H_1 και εφαρμόζοντας το Λημμα 1.6.1 (ii) του παραρτήματος, παίρνουμε οτι ο περιορισμός του cb^- στην Imb δεν εξαρτάται από την επιλογή του b^-. Επομένως, ο περιορισμός του ενδομορφισμού του Θ^*, ${}^t cqcb^-$ στην Imb είναι ορισμένος χωρίς αμφιβολία. Έστω Σ το φάσμα του ενδομορφισμού αυτού, σημειώνουμε $\sigma_M = sup\Sigma$ και $\sigma_m = inf\Sigma$, το οποίο είναι θετικό ή μηδέν, σύμφωνα με το Λήμμα 1.6.1(iv) του παραρτήματος, (αφού κάθε τετραγωνική μορφή δίνει τιμές στον $\mathbb{R}_+$). Έχουμε λοιπόν:

Πρόταση 1.3.1

Υπό την υπόθεση H_1, μια ικανή συνθήκη ώστε ο κίνδυνος του ϕ να είναι πεπερασμένος στο (θ, σ) είναι: $\mathbb{E}_{\theta,\sigma}[th^2(t, 1/s^2)] < +\infty$.

Εάν επιπλέον, $rg(c) = rg(b)$ και ο περιορισμός του q στην Imc είναι θετικά ορισμένος τότε αυτή η συνθήκη είναι και αναγκαία.

Απόδειξη

Αν θεωρήσουμε $a = {}^t cqc$ έχουμε οτι ,

$$\forall x \in \Theta,\ a(x) = {}^t cqc(x) = {}^t cq\Big(c(x),\ .\ \Big) = q\Big(c(x),\ c(.)\Big),\ \text{άρα},\ \bar{a}(x) = \bar{q}\Big(c(x)\Big)\ .$$

Οπότε σύμφωνα με το Λήμμα 1.6.1 (iv) έχουμε :

$$\forall x \in \Theta\ ,\quad \sigma_m \bar{b}(x) \leq \bar{q}(c(x)) \leq \sigma_M \bar{b}(x) \tag{1.2}$$

Για $x = \phi_0(y)$, πολλαπλασιάζοντας και τα δύο μέλη αυτής της ανισότητας με $h^2(t, 1/s^2)$ παίρνουμε :

$$\sigma_m t h^2(t, 1/s^2) \leq \bar{q}\Big(c(\phi_0(y))\Big) h^2(t, 1/s^2) \leq \sigma_M t h^2(t, 1/s^2)$$

και ολοκληρώνοντας αναφορικά με τον $N(\theta, \sigma^2 v)$, έχουμε :

$$\sigma_m \mathbb{E}_{\theta,\sigma}[t h^2(t, 1/s^2)] \leq \mathbb{E}_{\theta,\sigma}[\bar{q}(c(\phi_0(y)))h^2(t, 1/s^2)] \leq \sigma_M \mathbb{E}_{\theta,\sigma}[t h^2(t, 1/s^2)] \tag{1.3}$$

Αν $\mathbb{E}_{\theta,\sigma}[t h^2(t, 1/s^2)] < +\infty$, από τη σχέση (1.3) προκύπτει οτι

$$\mathbb{E}_{\theta,\sigma}\Big[\bar{q}\Big(c(\phi_0(y))\Big) h^2(t, 1/s^2)\Big] < +\infty$$

Όμως,

$$R_\phi(\theta,\sigma) = R_{\phi_0}(\theta,\sigma) + \sigma^{-2}\mathbb{E}_{\theta,\sigma}\Big[\bar{q}\Big(\phi(y) - \phi_0(y)\Big)\Big] - 2\sigma^{-2}\mathbb{E}_{\theta,\sigma}\Big[q\Big(\phi_0(y) - \phi(y), \phi_0(y) - \theta\Big)\Big]$$

όπου, εφαρμόζοντας την ανισότητα του *Schwarz* στον τελευταίο όρο,

$$\begin{aligned} \mathbb{E}_{\theta,\sigma}\Big[q\Big(\phi_0(y) - \phi(y),\ \phi_0(y) - \theta\Big)\Big] &\leq \mathbb{E}_{\theta,\sigma}\Big[\bar{q}^{\ 1/2}(\phi_0(y) - \phi(y))\ \bar{q}^{\ 1/2}(\phi_0(y) - \theta)\Big] \\ &\leq \mathbb{E}^{1/2}_{\theta,\sigma}\Big[\bar{q}(\phi_0(y) - \phi(y))\Big]\ \mathbb{E}^{1/2}_{\theta,\sigma}\Big[\bar{q}(\phi_0(y) - \theta)\Big] \end{aligned}$$

Έτσι, αν $\mathbb{E}_{\theta,\sigma}[\ \bar{q}(\phi_0(y) - \phi(y))\ ] = \mathbb{E}_{\theta,\sigma}[\ h^2(t, 1/s^2)\bar{q}(\ c(\ \phi_0(y)\ )\ )] < +\infty$, $R_\phi(\theta,\sigma) < +\infty$ και αντιστρόφως , αφού λόγω της παραπάνω ανισότητας ,

$$R_\phi(\theta,\sigma) \geq \left(R^{1/2}_{\phi_0}(\theta,\sigma) - \sigma^{-1}\ \mathbb{E}^{1/2}_{\theta,\sigma}\Big[\bar{q}(\phi_0(y) - \phi(y))\Big]\right)^2.$$

Λοιπόν , ο κίνδυνος του εκτιμητή ϕ είναι πεπερασμένος στο (θ, σ) αν και μόνον αν

$$\mathbb{E}_{\theta,\sigma}\Big[h^2(t, 1/s^2)\bar{q}(c(\ \phi_0(y)\ )\ )\Big] < +\infty$$

(συνθήκη αναφερόμενη για παράδειγμα από τους *Cellier* και *Fourdrinier*[1985]), απ᾽ όπου προκύπτει το πρώτο μέρος της πρότασης.

Αν επιπλέον $rg(c) = rg(b)$ και ο περιορισμός του q στην Imc είναι θετικά ορισμένος τότε η προηγούμενη συνθήκη είναι και αναγκαία. Πράγματι, λόγω της ικανής και αναγκαίας συνθήκης για το πεπερασμένο του κινδύνου που απεδείχθει προηγουμένως και της σχέσης (1.3) αρκεί να δείξουμε οτι το σ_m είναι γνήσια θετικό (δηλ. διάφορο του μηδενός). Όμως από την υπόθεση H_1 και επειδή $rg(c) = rg(b)$ θα ισχύει οτι $Ker\, b = Kerc$. Τότε:

$$\forall x \notin Ker\, b\ \ (\Leftrightarrow x \notin Kerc\ )\ ,\quad c(x) \neq 0$$

Άρα:

$$\forall x \notin Ker\, b\ ,\ \ \bar{q}\Big(c(x)\Big) \neq 0$$

αφού η q είναι θετικά ορισμένη στην Imc. Συνεπώς,

$$\sigma_m = \inf_{x \notin Ker\, b} \frac{\bar{q}(c(x))}{\bar{b}(x)} > 0$$

□

Παρατήρηση 1.3.2

Αν $\sigma_M = 0$ τότε όλες οι ιδιοτιμές του ενδομορφισμού ${}^t cqcb^-$ είναι μηδέν και λοιπόν ${}^t cqc$ είναι ταυτοτικά μηδέν. Επομένως, για κάθε (θ, y), $\bar{q}(\ \phi_0(y) - \theta\ ) = \bar{q}(\ \phi(y) - \theta\ )$. Δηλαδή, οι ϕ και ϕ_0 έχουν τον ίδιο κίνδυνο. Αυτό δικαιολογεί την υπόθεση:

ΥΠΟΘΕΣΗ $\mathbf{H_2}$: Ο περιορισμός του q στην Imc δεν είναι μηδέν .

Υπό την υπόθεση H_1, η H_2 ισοδυναμεί με την $\sigma_M \neq 0$, το οποίο προκύπτει άμεσα αφού:

$$\sigma_M = \sup_{x \notin Ker\, b} \frac{\bar{q}(c(x))}{\bar{b}(x)}$$

1.3.2

Πρόταση 1.3.2

Αν η υπόθεση H_1 ικανοποιείται και αν ο ϕ είναι ομαλός , ο κίνδυνος $R_\phi(\theta, \sigma)$ γράφεται :

$$R_\phi(\theta,\sigma) = R_{\phi_0}(\theta,\sigma) - 2\ \mathbb{E}_{\theta,\sigma}\Big[\ h(t,1/s^2)\ tr(v_\Theta qc)\ +\ 2\ h_1'(t,1/s^2)\ \overline{bv_\Theta qc}(\ \phi_0(y)\ )\ \Big]$$

$$+\ \mathbb{E}_{\theta,\sigma}\Big[\ \bar{q}(\ c(\phi_0(y))\ )\Big\{-4h(t,1/s^2)h_2'(t,1/s^2)\frac{1}{(n-k)s^4} + \frac{n-k-2}{(n-k)s^2}\ h^2(t,1/s^2)\ \Big\}\ \Big]$$

Παρατήρηση 1.3.3

Αυτό το αποτέλεσμα, γενίκευση του Θεωρήματος 1 των $Efron$ και $Morris$[1976], δίνει έναν *αμερόληπτο εκτιμητή* της διαφοράς των κινδύνων των εκτιμητών ϕ_0 και ϕ . Συγκεκριμένα, ο αμερόληπτος εκτιμητής αυτής της διαφοράς είναι η στατιστική συνάρτηση:

$$\begin{aligned}\Psi(y) \ &=\ 2\ h(t,1/s^2)\ tr(v_\Theta qc)\ +\ 2\ h_1'(t,1/s^2)\ \overline{bv_\Theta qc}(\phi_0(y)) \\ &\quad - \bar{q}(\ c(\phi_0(y))\ )\Big\{-4h(t,1/s^2)h_2'(t,1/s^2)\frac{1}{(n-k)s^4}\ +\ \frac{n-k-2}{(n-k)s^2}h^2(t,1/s^2)\Big\}\end{aligned}$$

Απόδειξη

Θεωρώντας τη βάση του E που εισαγάγαμε στην §1.2.3, μπορούμε να γράψουμε:

$$(v_\Theta qc)(\phi_0(y)) = \sum_{i=1}^{k}(d_i y_i + \beta_i)e_i$$

όπου για κάθε i, β_i είναι συνάρτηση των συντελεστών του $\phi_0(y)$ εκτός του y_i και d_i είναι μια σταθερά. Πράγματι,

$$\begin{aligned}
(v_\Theta qc)\Big(\phi_0(y)\Big) &= (v_\Theta q)\Big(\sum_{l=1}^k\sum_{j=1}^k c_{lj}y_je_l\Big) \;=\; v_\Theta\Big(\sum_{l=1}^k\sum_{j=1}^k c_{lj}y_jq(e_l)\Big)\\
&= v_\Theta\Big(\sum_{l=1}^k\sum_{j=1}^k c_{lj}y_j\sum_{i=1}^k q_{il}e_i^*\Big) \;=\; \sum_{l=1}^k\sum_{j=1}^k c_{lj}y_j\sum_{i=1}^k q_{il}v_\Theta(e_i^*)\\
&= \sum_{l=1}^k\;\sum_{j=1}^k c_{lj}y_j\sum_{i=1}^k q_{il}e_i \;=\; \sum_{i=1}^k e_i\sum_{j=1}^k\sum_{l=1}^k c_{lj}y_jq_{il}\\
&= \sum_{i=1}^k e_i\Big(\sum_{\substack{j=1\\ j\neq i}}^k\sum_{l=1}^k c_{lj}y_jq_{il} \;+\; \sum_{l=1}^k c_{li}y_iq_{il}\Big)\\
&= \sum_{i=1}^k (d_iy_i+\beta_i)
\end{aligned}$$

όπου :

$$\beta_i \;=\; \sum_{\substack{j=1\\ j\neq i}}^k\sum_{l=1}^k c_{lj}y_jq_{il} \qquad \text{και} \qquad d_i=\sum_{l=1}^k c_{li}q_{il}$$

Ορίζουμε $\Delta_\phi(\theta,\sigma)=R_\phi(\theta,\sigma)-R_{\phi_0}(\theta,\sigma)$. Οπότε προκύπτει:

$$\begin{aligned}
\Delta_\phi(\theta,\sigma) &= -\frac{2}{\sigma^2}\,\mathbb{E}_{\theta,\sigma}\Big[h(t,1/s^2)\;q\Big(\phi_0(y)-\theta,c(\phi_0(y))\Big)\Big] + \frac{1}{\sigma^2}\,\mathbb{E}_{\theta,\sigma}\Big[h^2(t,1/s^2)\;\bar q\Big(c(\phi_0(y))\Big)\Big]\\
&= -2A_\phi(\theta,\sigma) \;+\; B_\phi(\theta,\sigma)
\end{aligned}$$

Έχουμε (όπως προηγουμένως για τον υπολογισμό του $(v_\Theta qc)(\phi_0(y))$):

$$\begin{aligned}
q\Big(\phi_0(y)-\theta\;,\;c(\phi_0(y))\Big) &= q\Big(c(\phi_0(y))\Big)(\phi_0(y)-\theta)\\
&= \Big(\sum_{i=1}^k(d_iy_i+\beta_i)e_i^*\Big)\Big(\sum_{j=1}^k(y_j-\theta_j)e_j\Big)\\
&= \sum_{i=1}^k(d_iy_i+\beta_i)(y_i-\theta_i)
\end{aligned}$$

Άρα :

$$A_\phi(\theta,\sigma) \;=\; \frac{1}{\sigma^2}\,\mathbb{E}_{\theta,\sigma}\Big[\;h(t,1/s^2)\sum_{i=1}^k(y_i-\theta_i)(d_iy_i+\beta_i)\;\Big]$$

Πρέπει να σημειώσουμε οτι $\mathbb{E}_{\theta,\sigma}[\;h(t,1/s^2)(y_i-\theta_i)(d_iy_i+\beta_i)]$ είναι πεπερασμένη για κάθε i $(1\le i\le k)$. Πράγματι, από την ανισότητα του *Schwarz*, παίρνουμε:

$$\mathbb{E}_{\theta,\sigma}\Big[|\;h(t,1/s^2)(y_i-\theta_i)(d_iy_i+\beta_i)|\Big] \le \Bigg(\mathbb{E}_{\theta,\sigma}\Big[h^2(t,1/s^2)(d_iy_i+\beta_i)^2\Big]\;\underbrace{\mathbb{E}_{\theta,\sigma}\Big[(y_i-\theta_i)^2\Big]}_{=\sigma^2}\Bigg)^{1/2}$$

και

$$\begin{aligned}\mathbb{E}_{\theta,\sigma}\Big[h^2(t,1/s^2)(d_iy_i+\beta_i)^2\Big] &\leq \mathbb{E}_{\theta,\sigma}\Big[h^2(t,1/s^2)\sum_{j=1}^{k}(d_jy_j+\beta_j)^2\Big]\\ &\leq \mathbb{E}_{\theta,\sigma}\Big[h^2(t,1/s^2)\;\bar{v}_\Theta(\;qc(\phi_0(y))\;)\Big]\end{aligned}$$

Βλέπουμε λοιπόν, οτι αυτή η μέση τιμή είναι πεπερασμένη από μια ανισότητα την οποία μπορούμε να βγάλουμε ακριβώς όπως την (1.3), για $a = {}^tc^tqv_\Theta qc$ και χρησιμοποιώντας οτι η $\mathbb{E}_{\theta,\sigma}[\;t\,h^2(t,1/s^2)\;]$ είναι πεπερασμένη από την υπόθεση ομαλότητας του ϕ .

Μπορούμε επίσης να γράψουμε, για $1 \leq i \leq k$,

$$\mathbb{E}_{\theta,\sigma}\Big[h(t,1/s^2)\,(y_i-\theta_i)\,(d_iy_i+\beta_i)\Big]$$

$$= \int_{\mathbb{R}^{n-1}}\int_{-\infty}^{+\infty} h(t,1/s^2)(y_i-\theta_i)(d_iy_i+\beta_i)\frac{1}{\sqrt{2\pi}\,\sigma}\;e^{-(y_i-\theta_i)^2/2\sigma^2}\;dy_i\;\; N_{n-1}(\theta^i,\sigma^2I_{n-1};dy^i)$$

όπου, $\theta^i = (\;\theta_1,\;\ldots\;,\theta_{i-1},\theta_{i+1},\;\ldots\;,\theta_k,0,\;\ldots\;,0\;)$ και $y^i = (\;y_1,\;\ldots\;,y_{i-1},y_{i+1},\;\ldots\;,y_n\;)$.

Η $\mathbb{E}_{\theta,\sigma}[\;h(t,1/s^2)]$ είναι πεπερασμένη. Πράγματι, εφαρμόζοντας την ανισότητα του $Schwarz$, παίρνουμε:

$$\mathbb{E}_{\theta,\sigma}[\;h(t,1/s^2)\;] = \mathbb{E}_{\theta,\sigma}[\;t^{1/2}\;h(t,1/s^2)\;t^{-1/2}\;] \leq \Big(\mathbb{E}_{\theta,\sigma}[\;t\;h^2(t,1/s^2)\;]\;\;\mathbb{E}_{\theta,\sigma}[1/t]\Big)^{1/2}$$

Όμως, η $\mathbb{E}_{\theta,\sigma}[th^2(t,1/s^2)]$ είναι πεπερασμένη από την υπόθεση ομαλότητας του ϕ και η $\mathbb{E}_{\theta,\sigma}[1/t]$ είναι φραγμένη υπό μια πολλαπλασιαστική σταθερά, από τη ροπή τάξης -1 ενός μη κεντρικού $\mathcal{X}^2$- νόμου, με $rg(b)$ βαθμούς ελευθερίας. Πράγματι,

$$\begin{aligned}\mathbb{E}_{\theta,\sigma}[1/t] &= \int_{\mathbb{R}^n}\Big(\sum_{i=1}^{m}b_iy_i^2\Big)^{-1}N_n(\;\theta,\;\sigma^2I_n\;;\;dy_1,\;\ldots\;,dy_n)\\ &\leq \Big(\min_{1\leq i\leq m}b_i\Big)^{-1}\int_{\mathbb{R}^n}\Big(\sum_{i=1}^{m}y_i^2\Big)^{-1}N_n(\;\theta,\;\sigma^2I_n\;;\;dy_1,\;\ldots\;,dy_n)\end{aligned}$$

Επειδή $\theta = {}^t(\theta_1,\;\ldots\;,\theta_k,0,\;\ldots\;,0)$ και $m \leq k$,

$$\begin{aligned}&\int_{\mathbb{R}^n}\Big(\sum_{i=1}^{m}y_i^2\Big)^{-1}N_n(\theta,\sigma^2I_n;dy_1,\ldots,dy_n)\\ &= \int_{\mathbb{R}^m}\Big(\sum_{i=1}^{m}y_i^2\Big)^{-1}N_m\Big({}^t(\theta_1,\ldots,\theta_m),\sigma^2I_m;dy_1,\ldots,dy_m\Big)\\ &= \frac{1}{\sigma^2}\int_{\mathbb{R}^m}\Big(\sum_{i=1}^{m}z_i^2\Big)^{-1}N_m\Big(\frac{1}{\sigma}\;{}^t(\theta_1,\ldots,\theta_m),I_m;dz_1,\ldots,dz_m\Big)\\ &= \frac{1}{\sigma^2}\int_{\mathbb{R}_+}u^{-1}\;\mathcal{X}_m^2\Big(\frac{1}{\sigma^2}\;\sum_{i=1}^{m}\theta_i^2\;;\;du\Big)\end{aligned}$$

το οποίο είναι πεπερασμένο υποθέτοντας $m(=rg(b)\;) \geq 3$ (βλ. Πόρισμα 1, Παράρτημα).

Από το πεπερασμένο αυτής της μέσης τιμής για την h και τους περιορισμούς που επιβάλλονται από την ομαλότητα του ϕ, μπορούμε να επαληθεύσουμε την υπόθεση του Λήμματος του *Stein* (Λήμμα 1.6.2). Πράγματι, για $1 \leq i \leq k$,σημειώνοντας με f_i την πυκνότητα του μονοδιάστατου κανονικού νόμου με μέση τιμή θ_i και διασπορά σ^2, έχουμε :

$$
\begin{aligned}
&\int_{-\infty}^{+\infty} \Big| \frac{\partial}{\partial y_i}[h(t,1/s^2)\ (d_i y_i + \beta_i)] \Big| f_i(y_i)\, dy_i \\
&= \int_{-\infty}^{+\infty} \mid h_1'(t,1/s^2)\ 2b_i y_i\ (d_i y_i + \beta_i)\ +\ d_i h(t,1/s^2) \mid f_i(y_i)\, dy_i \\
&\leq\ 2 \mid d_i \mid \int_{-\infty}^{+\infty} b_i y_i^2 \mid h_1'(t,1/s^2) \mid f_i(y_i)\, dy_i +\ \mid \beta_i \mid \int_{-\infty}^{+\infty} \mid h_1'(t,1/s^2)\ 2b_i y_i \mid f_i(y_i)\, dy_i \\
&\qquad\qquad +\ \mid d_i \mid \int_{-\infty}^{+\infty} \mid h(t,1/s^2) \mid f_i(y_i)\, dy_i \\
&\leq\ 2 \mid d_i \mid \int_{-\infty}^{+\infty} \mid t\ h_1'(t,1/s^2) \mid f_i(y_i)\, dy_i +\ \mid \beta_i \mid \int_{-\infty}^{+\infty} \mid h_1'(t,1/s^2)\ 2b_i y_i \mid f_i(y_i)\, dy_i \\
&\qquad\qquad +\ \mid d_i \mid \int_{-\infty}^{+\infty} \mid h(t,1/s^2) \mid f_i(y_i)\, dy_i
\end{aligned}
$$

Ο πρώτος και ο τελευταίος όρος του παραπάνω αθροίσματος είναι πεπερασμένοι σχεδόν παντού αφού τα αντίστοιχα ολοκληρώματα στον E είναι πεπερασμένα. Πράγματι, καθένα από αυτά είναι ένα εσωτερικό ολοκλήρωμα των ροπών $\mathbb{E}_{\theta,\sigma}[\mid t\ h_1'(t,1/s^2)\ \mid]$ και $\mathbb{E}_{\theta,\sigma}[h(t,1/s^2)]$ αντιστοίχως, που είναι πεπερασμένες. Μένει επομένως να ελένξουμε αν ο μεσαίος όρος είναι πεπερασμένος.

Έχουμε :

$$
\begin{aligned}
&\int_{-\infty}^{+\infty} \mid h_1'(t,1/s^2)\ 2b_i y_i \mid f_i(y_i)\, dy_i \\
&\leq\ 2\int_{-\infty}^{-1} \mid h_1'(t,1/s^2)\ b_i y_i^2 \mid f_i(y_i)\, dy_i\ +\ 2\int_{1}^{+\infty} \mid h_1'(t,1/s^2)\ b_i y_i^2 \mid f_i(y_i)\, dy_i \\
&\qquad\qquad +\int_{-1}^{1} \mid h_1'(t,1/s^2)\ 2b_i y_i \mid f_i(y_i)\, dy_i \\
&\leq\ 2\int_{-\infty}^{+\infty} \mid h_1'(t,1/s^2)\ b_i y_i^2 \mid f_i(y_i)\, dy_i\ +\int_{-1}^{1} \mid h_1'(t,1/s^2)\ 2b_i y_i \mid f_i(y_i)\, dy_i \\
&\leq\ 2\int_{-\infty}^{+\infty} \mid t\ h_1'(t,1/s^2) \mid f_i(y_i)\, dy_i\ +\int_{-1}^{1} \mid h_1'(t,1/s^2)\ 2b_i y_i \mid f_i(y_i)\, dy_i
\end{aligned}
$$

Εδώ, ο πρώτος όρος είναι όμοιος με αυτόν που εμφανίστηκε προηγουμένως και είναι πεπερασμένος (όπως προηγούμενα). Για τον δεύτερο όρο, παρατηρούμε οτι:

$$
\int_{-1}^{1} h_1'(t,1/s^2)\ 2b_i y_i\ f_i(y_i)\, dy_i\ = \int_{-1}^{1} \Big(\ \frac{\partial}{\partial y_i} h(t,1/s^2)\ \Big)\ f_i(y_i)\, dy_i
$$

$$= \left[\ h(t,1/s^2)\ f_i(y_i)\ \right]_{-1}^{1} - \int_{-1}^{1} h(t,1/s^2)\frac{y_i-\theta_i}{\sigma^2}\ f_i(y_i)\,dy_i$$

το οποίο είναι πεπερασμένο αφού η h είναι συνεχής συνάρτηση. Άρα, τελικά,

$$\int_{-\infty}^{+\infty} \Big|\ \frac{\partial}{\partial y_i}\Big(h(t,1/s^2)\ (d_i y_i+\beta_i)\Big)\ \Big|\ f_i(y_i)\,dy_i < +\infty$$

και επομένως από το Λήμμα του $Stein$ παίρνουμε οτι, για $1 \le i \le k$,

$$\begin{aligned}
&\int_{-\infty}^{+\infty} h(t,1/s^2)\ (d_i y_i+\beta_i)\ \frac{y_i-\theta_i}{\sigma^2}\ f_i(y_i)\,dy_i \\
&= \int_{-\infty}^{+\infty} \frac{\partial}{\partial y_i}\Big[h(t,1/s^2)\ (d_i y_i+\beta_i)\Big]\ f_i(y_i)\,dy_i \\
&= \int_{-\infty}^{+\infty} \Big[h_1'(t,1/s^2)\ 2b_i y_i\ (d_i y_i+\beta_i)\ +\ d_i h(t,1/s^2)\Big]\ f_i(y_i)\,dy_i
\end{aligned}$$

Προκύπτει λοιπόν,

$$\begin{aligned}
A_\phi(\theta,\sigma) &= \frac{1}{\sigma^2}\ \mathbb{E}_{\theta,\sigma}\Big[h(t,1/s^2)\sum_{i=1}^{k}(y_i-\theta_i)\ (d_i y_i+\beta_i)\Big] \\
&= \sum_{i=1}^{k}\ \mathbb{E}_{\theta,\sigma}\Big[h(t,1/s^2)\ \frac{y_i-\theta_i}{\sigma^2}\ (d_i y_i+\beta_i)\Big] \\
&= \sum_{i=1}^{k}\ \mathbb{E}_{\theta,\sigma}\Big[h_1'(t,1/s^2)\ 2b_i y_i(d_i y_i+\beta_i) +\ d_i h(t,1/s^2)\Big] \\
&= \mathbb{E}_{\theta,\sigma}\Big[2h_1'(t,1/s^2)\Big(\sum_{i=1}^{m} b_i y_i(d_i y_i+\beta_i)\Big) + h(t,1/s^2)\Big(\sum_{i=1}^{k} d_i\Big)\Big] \\
&= \mathbb{E}_{\theta,\sigma}\Big[2h_1'(t,1/s^2)(\overline{bv_\Theta qc})(\phi_0(y)) +\ h(t,1/s^2)tr(v_\Theta qc)\Big]
\end{aligned}$$

Επίσης έχουμε :

$$\begin{aligned}
B_\phi(\theta,\sigma) &= \frac{1}{\sigma^2}\ \mathbb{E}_{\theta,\sigma}\Big[h^2(t,1/s^2)\ \bar{q}(\ c(\phi_0(y))\ )\Big] \\
&= \frac{1}{\sigma^2}\ \mathbb{E}_{\theta,\sigma}\Big[s^2\ \frac{1}{s^2}h^2(t,1/s^2)\ \bar{q}(\ c(\phi_0(y))\ )\Big] \\
&= \frac{1}{\sigma^2}\ \mathbb{E}_{\theta,\sigma}\Big[\Big(\sum_{i=k+1}^{n} y_i^2\Big)\frac{1}{(n-k)s^2}\ h^2(t,1/s^2)\ \bar{q}(\ c(\phi_0(y))\ )\Big] \\
&= \sum_{i=k+1}^{n}\ \mathbb{E}_{\theta,\sigma}\Big[\frac{1}{(n-k)s^2}\ h^2(t,1/s^2)\ y_i\ \frac{y_i}{\sigma^2}\ \bar{q}(\ c(\phi_0(y))\ )\Big]
\end{aligned}$$

Οι υποθέσεις ομαλότητας για τον ϕ , μας εξασφαλίζουν την απαραίτητη προϋπόθεση για την εφαρμογή του Λήμματος του $Stein$ στο εσωτερικό ολοκλήρωμα,

$$\int_{-\infty}^{+\infty} \frac{1}{(n-k)s^2}\ h^2(t,1/s^2)\ y_i\ \frac{y_i}{\sigma^2}\ f_i(y_i)\,dy_i$$

που εμφανίζεται στην παραπάνω έκφραση του $B_\phi(\theta,\sigma)$. Εδώ με f_i, για $k+1 \le i \le n$, σημειώνουμε την πυκνότητα του μονοδιάστατου κανονικού νόμου με μέσο μηδέν και διασπορά σ^2. Πράγματι, για $k+1 \le i \le n$, έχουμε :

$$\begin{aligned}
&\int_{-\infty}^{+\infty} \Big| \frac{\partial}{\partial y_i} \Big(h^2(t,1/s^2)\ \frac{y_i}{(n-k)s^2} \Big) \Big| f_i(y_i)\, dy_i \\
&= \int_{-\infty}^{+\infty} \Big| -\frac{4y_i^2}{n-k}\ h(t,1/s^2)\ h_2'(t,1/s^2)\ \frac{1}{(n-k)s^6} \\
&\qquad\qquad + h^2(t,1/s^2)\ \frac{1}{(n-k)s^2} \Big(1 - \frac{2y_i^2}{(n-k)s^2}\Big) \Big| f_i(y_i)\, dy_i \\
&\le \int_{-\infty}^{+\infty} \Big| \frac{4y_i^2}{n-k}\ h(t,1/s^2)\ h_2'(t,1/s^2)\ \frac{1}{(n-k)s^6} \Big| f_i(y_i)\, dy_i \\
&\quad + \int_{-\infty}^{+\infty} \Big| h^2(t,1/s^2)\ \frac{1}{(n-k)s^2} \Big(1 - \frac{2y_i^2}{(n-k)s^2}\Big) \Big| f_i(y_i)\, dy_i \\
&\le \int_{-\infty}^{+\infty} \Big| 4s^2\ \frac{1}{(n-k)s^6}\ h(t,1/s^2)\ h_2'(t,1/s^2) \Big| f_i(y_i)\, dy_i \\
&\quad + \int_{-\infty}^{+\infty} h^2(t,1/s^2)\ \frac{1}{(n-k)s^2} \Big(1 + \frac{2y_i^2}{(n-k)s^2}\Big) f_i(y_i)\, dy_i \\
&\le\ 4 \int_{-\infty}^{+\infty} \Big| \frac{1}{(n-k)s^4}\ h(t,1/s^2)\ h_2'(t,1/s^2) \Big| f_i(y_i)\, dy_i \\
&\quad + \int_{-\infty}^{+\infty} h^2(t,1/s^2)\ \frac{1}{(n-k)s^2} \Big(1 + \frac{2s^2}{s^2}\Big) f_i(y_i)\, dy_i \\
&\le\ 4 \int_{-\infty}^{+\infty} \Big| \frac{1}{(n-k)s^4}\ h(t,1/s^2)\ h_2'(t,1/s^2) \Big| f_i(y_i)\, dy_i \\
&\quad + 3 \int_{-\infty}^{+\infty} h^2(t,1/s^2)\ \frac{1}{(n-k)s^2}\ f_i(y_i)\, dy_i
\end{aligned}$$

Όμως από τις υποθέσεις ομαλότητας του ϕ, έχουμε,

$$\mathbb{E}_{\theta,\sigma}\Big[\Big|\ \frac{t}{s^4}\ h(t,1/s^2)\ h_2'(t,1/s^2)\ \Big|\Big] < +\infty$$

$$\text{δηλ.} \int_{\mathbb{R}^{n-1}} t\Big(\int_{-\infty}^{+\infty} \Big| \frac{1}{s^4}\ h(t,1/s^2)\ h_2'(t,1/s^2)\ \Big| f_i(y_i)\, dy_i \Big) N_{n-1}(\theta', \sigma^2 I_{n-1}\ ;\ dy^i) < +\infty$$

όπου $\theta' = {}^t(\theta_1, \ldots,\ \theta_k,\ 0, \ldots,\ 0) \in \mathbb{R}^{n-1}$ και $y^i = {}^t(y_1, \ldots,\ y_{i-1},\ y_{i+1}, \ldots,\ y_n)$. Επομένως, για $k+1 \le i \le n$,

$$\int_{-\infty}^{+\infty} \Big| \frac{1}{s^4}\ h(t,1/s^2)\ h_2'(t,1/s^2) \Big| f_i(y_i)\, dy_i < +\infty \ .$$

Επίσης ,

$$\mathbb{E}_{\theta,\sigma}\Big[\frac{t}{s^2}\ h^2(t,1/s^2)\Big] < +\infty$$

το οποίο σημαίνει οτι

$$\int_{\mathbb{R}^{n-1}} t\Big(\int_{-\infty}^{+\infty} \frac{1}{s^2}\; h^2(t,1/s^2)\; f_i(y_i)\,dy_i\Big) N_{n-1}(\theta', \sigma^2 I_{n-1}\;;\; dy^i) < +\infty$$

και λοιπόν, για $k+1 \leq i \leq n$,

$$\int_{-\infty}^{+\infty} \frac{1}{s^2}\; h^2(t,1/s^2)\; f_i(y_i)\,dy_i < +\infty \quad .$$

Άρα τελικά προκύπτει οτι :

$$\int_{-\infty}^{+\infty} \Big|\; \frac{\partial}{\partial y_i}\; (\; h^2(t,1/s^2)\; \frac{y_i}{(n-k)s^2}\;)\; \Big|\; f_i(y_i)\,dy_i < +\infty$$

και επομένως εφαρμόζοντας το Λήμμα του *Stein*, παίρνουμε για $k+1 \leq i \leq n$,

$$\begin{aligned}
&\int_{-\infty}^{+\infty} \frac{1}{(n-k)s^2}\; h^2(t,1/s^2)\; y_i\; \frac{y_i}{\sigma^2}\; f_i(y_i)\,dy_i \\
&= \int_{-\infty}^{+\infty} \frac{\partial}{\partial y_i} \Big(\; h^2(t,1/s^2)\; \frac{y_i}{(n-k)s^2}\; \Big) f_i(y_i)\,dy_i \\
&= \int_{-\infty}^{+\infty} \Big[- \frac{4y_i^2}{n-k}\; h(t,1/s^2)\; h_2'(t,1/s^2)\; \frac{1}{(n-k)s^6} \\
&\qquad + h^2(t,1/s^2)\; \frac{1}{(n-k)s^2}\Big(1 \;-\; \frac{2y_i^2}{(n-k)s^2}\;\Big)\;\Big] f_i(y_i)\,dy_i
\end{aligned}$$

Απ΄ όπου , σύμφωνα με τα προηγούμενα , έχουμε :

$$\begin{aligned}
B_\phi(\theta,\sigma) \;&=\; \mathbb{E}_{\theta,\sigma}\Big[\bar{q}(\; c(\phi_0(y))\;)\Big\{h^2(t,1/s^2)\; \frac{1}{(n-k)s^2}\; \Big(n-k\; -\; \frac{2\;\sum_{i=k+1}^n y_i^2}{(n-k)s^2}\Big) \\
&\qquad\qquad - 4\; \frac{\sum_{i=k+1}^n y_i^2}{n-k}\; h(t,1/s^2)\; h_2'(t,1/s^2)\; \frac{1}{(n-k)s^6}\Big\}\Big] \\
&=\; \mathbb{E}_{\theta,\sigma}\Big[\bar{q}(\; c(\phi_0(y))\;)\Big\{h^2(t,1/s^2)\; \frac{n-k-2}{(n-k)s^2} \\
&\qquad\qquad - 4\; h(t,1/s^2)\; h_2'(t,1/s^2)\; \frac{1}{(n-k)s^4}\Big\}\Big]
\end{aligned}$$

Συνεπώς,

$$\begin{aligned}
\Delta_\phi(\theta,\sigma) \;&=\; -\,2\;\mathbb{E}_{\theta,\sigma}\Big[h(t,1/s^2)\; tr(v_\Theta qc)\; +\; 2\; h_1'(t,1/s^2)\; \overline{bv_\Theta qc}(\; \phi_0(y)\;)\Big] \\
&\quad + \mathbb{E}_{\theta,\sigma}\Big[\bar{q}(\; c(\phi_0(y))\;)\Big\{h^2(t,1/s^2)\; \frac{n-k-2}{(n-k)s^2}\; -\; 4\; h(t,1/s^2)\; h_2'(t,1/s^2)\; \frac{1}{(n-k)s^4}\Big\}\Big]
\end{aligned}$$

□

1.3.3

Αν h η συνάρτηση συρρίκνωσης, θέτουμε:

$$r(x,z) \ = \ xzh(x,z)$$

Οι εκτιμητές με συρρικνωτή, γράφονται τότε:

$$\phi(y) \ = \ \Big(id_\Theta \ - \ \frac{s^2\, r(t,1/s^2)}{t}\ c\Big)\phi_0(y) \tag{1.4}$$

αυτή η γραφή είναι πράγματι κλασσική ($Baranchik$[1964], $Bock$[1975] , $Efron$ και $Morris$[1976], $Strawderman$[1973]). Συμπεραίνουμε λοιπόν, από την Πρόταση 1.3.2, το ακόλουθο πόρισμα:

Πόρισμα 1.3.1

Αν ικανοποιείται η υπόθεση H_1 και αν ο ϕ είναι ομαλός, ο κίνδυνος του παραπάνω εκτιμητή είναι:

$$\begin{aligned} R_\phi(\theta,\sigma) \ = \ R_{\phi_0}(\theta,\sigma) \ - \ 2\,\mathbb{E}_{\theta,\sigma}\Bigg[\frac{s^2}{t}\Big\{r(t,1/s^2)\ tr(v_\Theta qc) \\ + \ 2\overline{bv_\Theta qc}(\ \phi_0(y)\ )\Big(r_1'(t,1/s^2) - \frac{r(t,1/s^2)}{t}\Big)\Big\}\Bigg] \\ + \ \mathbb{E}_{\theta,\sigma}\Bigg[\frac{\bar{q}(\ c(\phi_0(y))\ )}{(n-k)t^2}\Big\{s^2\ r^2(t,1/s^2)\ (n-k+2)\ -4\ r(t,1/s^2)\ r_2'(t,1/s^2)\Big\}\Bigg] \end{aligned}$$

Πράγματι αυτό είναι άμεση συνέπεια της Προτάσεως 1.3.2, αν θέσουμε:

$$h_1'(t,1/s^2) = \frac{\partial}{\partial t}h(t,1/s^2) = \frac{\partial}{\partial t}\Big(\frac{s^2}{t}\ r(t,1/s^2)\Big) = -\frac{s^2}{t}\ r(t,1/s^2) + \frac{s^2}{t}\ r_1'(t,1/s^2)$$

και

$$h_2'(t,1/s^2) = \frac{\partial}{\partial(1/s^2)}h(t,1/s^2) = \frac{\partial}{\partial(1/s^2)}\Big(\frac{s^2}{t}\ r(t,1/s^2)\Big) = -\frac{s^4}{t}\ r(t,1/s^2) + \frac{s^2}{t}\ r_2'(t,1/s^2).$$

1.4 ΙΚΑΝΗ ΣΥΝΘΗΚΗ ΚΥΡΙΑΡΧΙΑΣ

1.4.1

Έστω $a = bv_\Theta qc$. Επειδή (λόγω της H_1) $Ker\, b \subset Ker\, a$, σύμφωνα με το Λήμμα 1.6.1 του παραρτήματος (Παρατήρηση 1.6.10) το φάσμα του ab^- συμπίπτει με αυτό του ενδομορφισμού $v_\Theta qc$. Έστω Λ το φάσμα του περιορισμού του ab^- στην Imb (ανεξάρτητο της επιλογής του b^- σύμφωνα με το Λήμμα 1.6.1(ii)). Αυτό είναι επίσης το σύνολο το οποίο λαμβάνεται παραλείποντας $(k - rg(b))-$μηδενικές ιδιοτιμές από το φάσμα του $v_\Theta qc$. Έστω:

$$\lambda_m = inf \Lambda \qquad \text{και} \qquad \lambda_M = sup\ \Lambda$$

Έχουμε τότε, συμφωνα με το Λήμμα 1.6.1(iv), για κάθε $y \in E$ τέτοιο ώστε $t = \bar{b}(\phi_0(y)) \neq 0$,

$$\lambda_m \leq \frac{\overline{bv_\Theta qc}\ (\ \phi_0(y)\ )}{t} \leq \lambda_M \tag{1.5}$$

Υπό τις υποθέσεις H_1 και H_2 , ορίζουμε :

$$R_0 = 2\ \frac{tr(v_\Theta qc)\ - 2\ \lambda_M}{\sigma_M}\ \frac{n-k}{n-k+2}$$

Αν $\lambda_M \geq 0$, για να είναι το R_0 θετικό, είναι αναγκαία η συνθήκη $rg(b) \geq 3$, επειδή $tr(v_\Theta qc) - 2\lambda_M \leq (rg(b)-2)\lambda_M$.

Σημειώνοντας με: $r_i'^{+} = sup(0, r_i')$ και $r_i'^{-} = sup(0, -r_i')$ $(i = 1, 2)$ προκύπτει η παρακάτω πρόταση:

Πρόταση 1.4.3

Αν οι υποθέσεις H_1 και H_2 ικανοποιούνται , αν ο ϕ είναι ομαλός και για κάθε (t, u) ,

$$\begin{aligned} 0 \ \leq \ & \sigma_M\ \frac{n-k+2}{n-k}\ r(t,u)\Big[R_0 - r(t,u)\Big] \\ & + 4\Big[t\Big(\lambda_m r_1'^{+}(t,u) - \lambda_M r_1'^{-}(t,u)\Big)\ +\ \frac{r(t,u)u}{n-k}\Big(\sigma_m r_2'^{+}(t,u) - \sigma_M r_2'^{-}(t,u)\Big)\Big] \end{aligned} \tag{1.6}$$

τότε ο ϕ κυριαρχεί τον ϕ_0 .

Απόδειξη

Σύμφωνα με την σχέση (1.5), πολλαπλασιάζοντας με $r_1'(t, 1/s^2)$, έχουμε:

$$\frac{r_1'(t,1/s^2)\ \overline{bv_\Theta qc}(\ \phi_0(y)\ )}{t} \geq \lambda_m\ r_1'^{+}(t,1/s^2) - \lambda_M\ r_1'^{-}(t,1/s^2)$$

και όμοια από την σχέση (1.2) προκύπτει :

$$\frac{r_2'(t,1/s^2)\ \bar{q}(\ \phi_0(y)\ )}{t} \geq \sigma_m\ r_2'^{+}(t,1/s^2) - \sigma_M\ r_2'^{-}(t,1/s^2)$$

Από το πόρισμα 1.3.1, θέτοντας $u = 1/s^2$, έχουμε οτι:

$$\begin{aligned} & R_{\phi_0}(\theta,\sigma)\ -\ R_\phi(\theta,\sigma) \\ = \ & \mathbb{E}_{\theta,\sigma}\Bigg[\frac{1}{tu}\Big\{2\ r(t,u)\ tr(v_\Theta qc) +\ 4\ \overline{bv_\Theta qc}(\ \phi_0(y)\ )\Big(r_1'(t,u)\ -\ \frac{r(t,u)}{t}\Big) \\ & \qquad -\ \frac{\bar{q}(\ c(\phi_0(y))\ )}{(n-k)t}\ r^2(t,u)\ (n-k+2)\ +\ 4\ \frac{\bar{q}(\ c(\phi_0(y))\ )\ u}{(n-k)t}\ r(t,u)\ r_2'(t,u)\Big\}\Bigg] \\ \geq \ & \mathbb{E}_{\theta,\sigma}\Bigg[\frac{1}{tu}\ \Big\{\sigma_M\ \frac{n-k+2}{n-k}\ R_0\ r(t,u)\ +\ 4\lambda_M r(t,u)\ +\ 4t\Big(\lambda_m r_1'^{+}(t,u)\ -\ \lambda_M r_1'^{-}(t,u)\Big) \\ & \qquad -\ 4\lambda_M r(t,u)\ -\ \frac{n-k+2}{n-k}\ \sigma_M r^2(t,u)\ +\ \frac{4u}{n-k}\ r(t,u)\Big(\sigma_m r_2'^{+}(t,u)\ -\ \sigma_M r_2'^{-}(t,u)\Big)\Big\}\Bigg] \\ \geq \ & 0\ , \ \text{λόγω της σχέσης (1.6).} \end{aligned}$$

□

Παρατήρηση 1.4.4
Αν η $r(t,u)$ είναι , για κάθε t , αύξουσα συνάρτηση του u και $0 \leq r(t,u) \leq R_0$, τότε :

(i) αν $\lambda_m \geq 0$ και $r(t,u)$ είναι , για κάθε u, αύξουσα συνάρτηση του t , τότε ο ϕ κυριαρχεί τον ϕ_0. Το αποτέλεσμα αυτό προκύπτει από την Πρόταση 1.4.3 , αφού η σχέση (1.6) επαληθεύεται:

$$\sigma_M \, \frac{n-k+2}{n-k} \, r(t,u)\Big[R_0 \; - \; r(t,u)\Big] + \; 4\Big[t \, \lambda_m \, r_1'(t,u) \; + \; \frac{r(t,u)\, u}{n-k} \, \sigma_m \, r_2'(t,u)\Big] \geq 0$$

(ii) αν $\lambda_M \leq 0$ και $r(t,u)$ είναι , για κάθε u, φθίνουσα συνάρτηση του t , τότε ο ϕ κυριαρχεί τον ϕ_0. Και αυτό προκύπτει από την Πρόταση 1.4.3, αφού η σχέση (1.6) επαληθεύεται:

$$\sigma_M \, \frac{n-k+2}{n-k} \, r(t,u)\Big[R_0 \; - \; r(t,u)\Big] + \; 4\Big[t \, \lambda_M \, r_1'(t,u) \; + \; \frac{r(t,u)\, u}{n-k} \, \sigma_m \, r_2'(t,u)\Big] \geq 0$$

(iii) αν $\lambda_m < 0$ και $r(t,u)$ είναι , για κάθε u, φθίνουσα συνάρτηση του t , τότε αν $\lambda_M \geq 0$ η σχέση (1.6) δεν οδηγεί σε συμπέρασμα , ενώ αν $\lambda_M < 0$ η σχέση αυτή επαληθεύεται και συνεπώς ο ϕ κυριαρχεί τον ϕ_0(από την Πρόταση 1.4.3) .

(iv) αν $\lambda_M > 0$ και $r(t,u)$ είναι, για κάθε u, αύξουσα συνάρτηση του t, τότε αν $\lambda_m < 0$ η σχέση (1.6) δεν οδηγεί σε συμπέρασμα, ενώ αν $\lambda_m \geq 0$ η σχέση αυτή επαληθεύεται και συνεπώς ο ϕ κυριαρχεί τον ϕ_0.
(άρα από τις (iii) και (iv) δεν προκύπτει λύση όταν $\lambda_m \lambda_M \leq 0$)

1.4.2 Μελέτη ιδιαιτέρων συναρτήσεων συρρίκνωσης

1.4.2(A) . Θεωρούμε την υπόθεση:

ΥΠΟΘΕΣΗ $\mathbf{H_3}$: Υπάρχει μια συνάρτηση g από τον $\mathbb{R}_+$ στον $\mathbb{R}_+$, τέτοια ώστε, για κάθε $(t,u) \in (\mathbb{R}_+)^2$, $r(t,u) = g(tu)$ και το $\Delta = \{x : g(x) = R_0\}$ είναι μέτρου $(Lebesgue)$ μηδέν. Επιπλέον $R_0 \lambda_m \lambda_M \neq 0$.

Αυτή η περίπτωση, όπου η συνάρτηση συρρίκνωσης εξαρτάται από μία μόνο μεταβλητή είναι αυτή που παρουσιάζεται συχνότερα στη βιβλιογραφία ($Baranchik$[1964], $Cellier$ και $Fourdrinier$ [1985], $Efron$ και $Morris$[1976], $James$ και $Stein$[1961]).

Υπό την υπόθεση H_3, ορίζουμε τις συναρτήσεις F_1 και F_2 από την παρακάτω σχέση:

$$\forall x \notin \; \Delta \; , \;\; F_i(x) = \frac{\lambda_i \; x^{\alpha_i} \; g(x) \; sgn(R_0 - g(x))}{R_0 \; |R_0 - g(x)|^{a_i}} \qquad (i = 1, 2)$$

με : $\lambda_1 = \lambda_m$, $\lambda_2 = \lambda_M$

$$a_1 \; = \; 1 \; + \; \frac{\sigma_m \; R_0}{(n-k) \; \lambda_m} \quad , \quad a_2 \; = \; 1 \; + \; \frac{\sigma_M \; R_0}{(n-k) \; \lambda_M}$$

και $\quad \alpha_i \; = \; \dfrac{(n-k+2) \; R_0 \; \sigma_M}{4 \; (n-k) \; \lambda_i} \quad , \quad (i = 1, 2)$

Πρόταση 1.4.4

Άν οι υποθέσεις H_1, H_2, H_3 ικανοποιούνται ,αν ο ϕ είναι ομαλός και αν για κάθε $x \notin \Delta$ έχουμε

$$g'(x) \geq 0 \;\; \textit{και} \;\; F_1'(x) \geq 0 \qquad \textit{ή} \qquad g'(x) \leq 0 \;\; \textit{και} \;\; F_2'(x) \geq 0 \quad ,$$

τότε ο ϕ κυριαρχεί τον ϕ_0.

Παρατήρηση 1.4.5

Συμπεραίνουμε έτσι, από τα προηγούμενα αποτελέσματα, μία γενίκευση του θεωρήματος 2 των $Efron$ και $Morris$[1976] και του 10.2.1.(b) των $Judge$ και $Bock$[1978].

Απόδειξη

Για $x \notin \Delta$ και τέτοιο ώστε $R_0 - g(x) > 0$ και $i = 1, 2$, έχουμε:

$$F_i(x) = \frac{\lambda_i \; x^{\alpha_i} \; g(x)}{R_0(\; R_0 - g(x) \;)^{a_i}}$$

απ' όπου παραγωγίζοντας παίρνουμε :

$$\begin{aligned} F_i'(x) \;\; &= \;\; \frac{\lambda_i R_0(\; R_0 - g(x) \;)^{a_i} \Big[\alpha_i x^{\alpha_i - 1} g(x) + x^{\alpha_i} g'(x) \Big] \; + \; \lambda_i x^{\alpha_i} g(x) R_0(\; R_0 - g(x) \;)^{a_i - 1} g'(x) a_i}{R_0^2 \; (\; R_0 - g(x) \;)^{2a_i}} \\ &= \;\; \frac{\lambda_i x^{\alpha_i - 1} \Big[\; \alpha_i \; g(x) \; (\; R_0 - g(x) \;) \; + \; R_0 \; x \; g'(x) \; + \; (a_i - 1) \; x \; g(x) \; g'(x) \; \Big]}{R_0 \; (\; R_0 - g(x) \;)^{a_i + 1}} \end{aligned}$$

Για $x \notin \Delta$ και τέτοιο ώστε $R_0 - g(x) < 0$ και $i = 1, 2$, έχουμε :

$$F_i(x) \;\; = \;\; - \; \frac{\lambda_i \; x^{\alpha_i} \; g(x)}{R_0(\; g(x) - R_0 \;)^{a_i}}$$

απ' όπου παραγωγίζοντας παίρνουμε :

$$\begin{aligned} F_i'(x) \;\; &= \;\; \frac{-\lambda_i R_0 (g(x) - R_0)^{a_i} \Big[\alpha_i x^{\alpha_i - 1} g(x) + x^{\alpha_i} g'(x) \Big] \; + \; \lambda_i x^{\alpha_i} g(x) R_0(\; g(x) - R_0 \;)^{a_i - 1} g'(x) a_i}{R_0^2 \; (\; g(x) - R_0 \;)^{2a_i}} \\ &= \;\; \frac{\lambda_i x^{\alpha_i - 1} \Big[\alpha_i g(x)(R_0 - g(x)) \; + \; R_0 x g'(x) \; + \; (a_i - 1) x g(x) g'(x) \Big]}{R_0 (g(x) - R_0)^{a_i + 1}} \end{aligned}$$

Άρα, τελικά, για κάθε $x \notin \Delta$ και $i = 1, 2$, έχουμε:

$$F_i'(x) \;\; = \;\; \frac{\lambda_i x^{\alpha_i - 1} \Big[\alpha_i g(x)(R_0 - g(x)) \; + \; R_0 x g'(x) \; + \; (a_i - 1) x g(x) g'(x) \Big]}{R_0 |R_0 - g(x)|^{a_i + 1}}$$

Εδώ, επειδή $r_1'(t, u) = u \; g'(tu)$ και $r_2'(t, u) = t \; g'(tu)$, η ανισότητα (1.6) γράφεται :

$$0 \;\leq\; \sigma_M \; \frac{n - k + 2}{n - k} \; g(tu) \Big(R_0 - g(tu) \Big) \; + \; 4 \Big(t \; \lambda_m \; u \; g'(tu) + \; \frac{g(tu) u}{n - k} \; \sigma_m t \; g'(tu) \Big), \text{αν } g'(tu) \geq 0$$

και

$$0 \ \leq \ \sigma_M \ \frac{n-k+2}{n-k} \ g(tu)\Big(R_0 - g(tu)\Big) \ + \ 4\Big(t \ \lambda_M u \ g'(tu) + \ \frac{g(tu)u}{n-k} \ \sigma_M t \ g'(tu)\Big), \text{αν } g'(tu) \leq 0$$

Όμως για $x \in (\mathbb{R}_+ - \Delta)$, έχουμε:

- αν $F_1'(x) \geq 0$, τότε

$$\frac{\lambda_m x^{\alpha_1 - 1}\Big[\alpha_1 g(x)(R_0 - g(x)) \ + \ R_0 x g'(x) \ + \ (a_1 - 1)x g(x) g'(x)\Big]}{R_0 |R_0 - g(x)|^{a_1+1}} \geq 0$$

απ' όπου, αντικαθιστώντας το α_1 και το a_1, προκύπτει:

$$\frac{\lambda_m}{R_0}\Big[\frac{(n-k+2)R_0\sigma_M}{4(n-k)\lambda_m} \ g(x)\Big(R_0 - g(x)\Big) \ + \ R_0 x g'(x) \ + \ \frac{\sigma_m R_0}{(n-k)\lambda_m} \ x g(x) g'(x)\Big] \geq 0$$

δηλαδή

$$\sigma_M \ \frac{n-k+2}{n-k} \ g(x)\Big(R_0 - g(x)\Big) \ + \ 4\Big(\lambda_m x g'(x) \ + \ \frac{4}{n-k} \ \sigma_m x g(x) g'(x)\Big) \geq 0$$

και ομοίως,

- αν $F_2'(x) \geq 0$, τότε

$$\frac{\lambda_M}{R_0}\Big[\frac{(n-k+2)R_0 \ \sigma_M}{4(n-k)\lambda_M} \ g(x)\Big(R_0 - g(x)\Big) \ + \ R_0 x g'(x) \ + \ \frac{\sigma_M R_0}{(n-k)\lambda_M} \ x g(x) g'(x)\Big] \geq 0$$

δηλαδή

$$\sigma_M \ \frac{n-k+2}{n-k} \ g(x)\Big(R_0 - g(x)\Big) \ + \ 4\Big(\lambda_M x g'(x) \ + \ \frac{4}{n-k} \ \sigma_M x g(x) g'(x)\Big) \geq 0$$

Από τις δύο τελευταίες σχέσεις για $x = tu$ και από την Πρόταση 1.4.3 προκύπτει το ζητούμενο
.

$\square$

1.4.2(B). Μερική περίπτωση

Στήν συνέχεια των *James* και *Stein*[1961] , μια περίπτωση που εμφανίζεται συχνά είναι εκείνη όπου η συνάρτηση r , ορισμένη στην παράγραφο 1.3.3. είναι σταθερή (βλέπε, για παράδειγμα, *Stein*[1981]). Λαμβάνουμε έτσι μία υπο-κλάση των εκτιμητών της μορφής (1.4),

$$\phi_\varrho(y) \ = \ \Big(\ id_\Theta \ - \ \varrho \ \frac{s^2}{t} \ c \ \Big)\phi_0(y) \ , \quad \varrho \in \mathbb{R}_+ \tag{1.7}$$

Όταν $b = v_\Theta^{-1}$, αυτοί οι εκτιμητές αποτελούν μια άμεση γενίκευση του "ιστορικού" εκτιμητή των *James* και *Stein*[1961] (για τον οποίο $c = id_\Theta$). Οι μεταβλητές t και s^2 είναι ανεξάρτητες και οι συνθήκες ομαλότητας ικανοποιούνται αν $\mathbb{E}_{\theta,\sigma}(1/t)$, $\mathbb{E}_{\theta,\sigma}(s^2)$ και $\mathbb{E}_{\theta,\sigma}(s^4)$ είναι πεπερασμένες. Έχουμε δεί στην απόδειξη της Πρότασης 1.3.2 οτι για $rg(b) \geq 3$, η πρώτη από τις συνθήκες αυτές ικανοποιείται. Οι άλλες δύο προκύπτουν από το οτι η $(n-k)s^2/\sigma^2$ ακολουθεί μια $\mathcal{X}^2$-κατανομή με $(n-k)$ βαθμούς ελευθερίας για την οποία υπάρχουν όλες οι ροπές τάξεως μεγαλύτερης ή ίσης του $-(n-k)/2$. Προκύπτει λοιπόν, από το Πόρισμα 1.3.1, οτι:

$$R_{\phi_\varrho}(\theta,\sigma) \;=\; R_{\phi_0}(\theta,\sigma) \;-\; 2\varrho\, \mathbb{E}_{\theta,\sigma}\Bigg[\; \frac{tr(v_\Theta qc)s^2}{t} \;-\; 2\; \frac{\overline{bv_\Theta qc}(\;\phi_0(y)\;)s^2}{t^2}\Bigg]$$

$$+\; \frac{n-k+2}{n-k}\; \varrho^2\; \mathbb{E}_{\theta,\sigma}\Bigg[\; \frac{\bar{q}\Big(c(\;\phi_0(y)\;)\Big)s^2}{t^2}\Bigg]$$

Σταθεροποώντας το (θ,σ), ο κίνδυνος $R_{\phi_\varrho}(\theta,\sigma)$ είναι συνάρτηση του ϱ (και μάλιστα πολυώνυμο δευτέρου βαθμού), φθίνει μέχρι το ϱ να πάρει την τιμή:

$$\varrho_m(\theta,\sigma) \;=\; \frac{n-k}{n-k+2}\; \frac{tr(v_\Theta qc)\; \mathbb{E}_{\theta,\sigma}\Big[s^2/t\Big] \;-\; 2\; \mathbb{E}_{\theta,\sigma}\Big[\overline{bv_\Theta qc}(\;\phi_0(y)\;)\; s^2/t^2\Big]}{\mathbb{E}_{\theta,\sigma}\Big[\bar{q}\Big(c(\;\phi_0(y)\;)\Big)s^2/t^2\Big]}$$

και αυξάνει όταν το ϱ παίρνει μεγαλύτερες τιμές από αυτήν.

Από την σχέση (1.2) για $x=\phi_0(y)$ έχουμε, $\sigma_m t \leq \bar{q}(\;c(\phi_0(y))\;) \leq \sigma_M t$. Χρησιμοποιώντας τη σχέση αυτή και την (1.5), έχουμε:

$$\varrho_m(\theta,\sigma) \geq \frac{n-k}{n-k+2}\; \frac{tr(v_\Theta qc)\; \mathbb{E}_{\theta,\sigma}[s^2/t] \;-\; 2\;\lambda_M\; \mathbb{E}_{\theta,\sigma}[s^2/t]}{\sigma_M\; \mathbb{E}_{\theta,\sigma}[s^2/t]} \;=\; \frac{R_0}{2}$$

και

$$\varrho_m(\theta,\sigma) \leq \frac{n-k}{n-k+2}\; \frac{tr(v_\Theta qc)\; \mathbb{E}_{\theta,\sigma}[s^2/t] \;-\; 2\;\lambda_m\; \mathbb{E}_{\theta,\sigma}[s^2/t]}{\sigma_m\; \mathbb{E}_{\theta,\sigma}[s^2/t]} \;=\; \frac{R_1}{2}$$

Επομένως,

$$\forall(\theta,\sigma)\;,\quad \frac{R_0}{2} \leq \varrho_m(\theta,\sigma) \leq \frac{R_1}{2}$$

όπου $R_1 \;=\; 2\; \dfrac{tr(v_\Theta qc) \;-\; 2\;\lambda_m}{\sigma_m}\; \dfrac{n-k}{n-k+2}\;,$ αν $\sigma_m \neq 0$ (διαφορετικά αντικαθιστούμε το σ_m με την πιο μικρή μη-μηδενική ιδιοτιμή του ${}^t cqcb^-$). Παρατηρούμε οτι, αν το R_1 είναι αρνητικό, οι βέλτιστοι (*optimals*) συντελεστές ϱ είναι αρνητικοί, το οποίο αποκλείεται. Υποθέτουμε, λοιπόν, στα παρακάτω ότι το R_1 είναι θετικό [μπορούμε να επανέλθουμε σε αυτό μεταβάλλοντας ενδεχομένως το c σε $(-c)$].

Πρόταση 1.4.5

Μεταξύ των εκτιμητών ϕ_ϱ, αυτοί για τους οποίους το ϱ περιορίζεται μεταξύ του $sup(0,R_0/2)$ και του $R_1/2$, σχηματίζουν μία πλήρη κλάση.

Αν $\varrho \geq R_1$, τότε ο ϕ_0 κυριαρχεί τον ϕ_ϱ. Όταν το $R_0 \geq 0$, ο ϕ_ϱ κυριαρχεί τον ϕ_0 για $\varrho \in [0,R_0]$.

Απόδειξη

Το πρώτο μέρος της παραπάνω πρότασης είναι προφανές από τον ορισμό της πλήρης κλάσης. (Αν $\mathcal{E}=\{\phi : E \to \Theta\}$ ο χώρος των εκτιμητών του θ και $\mathcal{C} \subseteq \mathcal{E}$ τότε η $\mathcal{C}$ είναι πλήρης αν $\forall\phi \notin \mathcal{C},\ \exists\phi' \in \mathcal{C}$ τέτοιος ώστε $R_{\phi'} < R_\phi$.)

Για το δεύτερο μέρος της πρότασης έχουμε : $\varrho \geq R_1 \Rightarrow \varrho \geq 2\varrho_m$. Όμως, για $\varrho = 2\varrho_m$ έχουμε $R_{\phi_\varrho}(\theta,\sigma) = R_{\phi_0}(\theta,\sigma)$ και για $\varrho \geq 2\varrho_m$, η $R_{\phi_\varrho}(\theta,\sigma)$ είναι αύξουσα συνάρτηση του ϱ. Άρα για $\varrho \geq R_1$, $R_{\phi_\varrho}(\theta,\sigma) \geq R_{\phi_0}(\theta,\sigma)$.

Όταν το $R_0 \geq 0$, από την Πρόταση 1.4.3 έχουμε οτι, αν

$$0 \;\leq\; \sigma_M\, \frac{n-k+2}{n-k}\, \varrho\,(R_0-\varrho) \;\; (\text{δηλαδή για } \varrho \;\leq\; R_0)$$

τότε ο ϕ κυριαρχεί τον ϕ_0.

$\square$

Παρατήρηση 1.4.6

Όταν $R_0R_1 \leq 0$, δεν μπορούμε να πούμε κάτι για την διαφορά των κινδύνων μεταξύ του ϕ_ϱ και του ϕ_0, για $\varrho \in [0, R_1/2]$. Για ορισμένες τιμές του (θ, σ), ο ϕ_0 μπορεί να έχει μικρότερο κίνδυνο από αυτόν του ϕ_ϱ, για οποιονδήποτε ϱ.

Παρατήρηση 1.4.7

Σε αντίθεση με τους $James$ και $Stein$[1961] και τον $Stein$[1981], η έκφραση του $\varrho_m(\theta,\sigma)$ δείχνει οτι δεν υπάρχει γενικά ένας εκτιμητής ϕ_ϱ, βέλτιστος ομοιομόρφως στο (θ,σ). (Ο $Stein$[1981] μελέτησε την ειδική περίπτωση όπου $b = \upsilon_\Theta^{-1}c^2$ με c σταθερά).

1.5 ΔΙΑΦΟΡΙΣΙΜΟΙ ΕΛΕΓΧΟΜΕΝΟΙ ΕΚΤΙΜΗΤΕΣ

1.5.1

Οι συνθήκες τις οποίες έχουμε καθορίσει στις προηγούμενες παραγράφους φαίνονται να είναι είτε πολύ γενικές και γι'αυτό δύσκολα εφαρμόσιμες (Πρόταση 1.4.3) είτε πολύ συγκεκριμένες (Προτάσεις 1.4.4 και 1.4.5) λόγω των περιοριστικών υποθέσεων επί της συνάρτησης συρρίκνωσης. Οι περιορισμοί που θα εισάγουμε παρακάτω είναι πιο ισχυροί από αυτούς της Πρότασης 1.4.3, αλλά παρουσιάζουν το πλεονέκτημα να χαρακτηρίζουν την μεταβολή της συνάρτησης συρρίκνωσης απαιτώντας μία ελλάτωση στο αλγεβρικό πλαίσιο του μοντέλου.

Επιβάλουμε πράγματι, στον ενδομορφισμό $\upsilon_\Theta qc$ να έχει όλες τις ιδιοτιμές του θετικές ή μηδέν. Δηλαδή, υποθέτουμε:

ΥΠΟΘΕΣΗ $\mathbf{H_4}$: $\lambda_m \geq 0$.

Σημειώνουμε οτι αυτή η υπόθεση επιβάλλει στον c να έχει όλες τις ιδιοτιμές του θετικές ή μηδέν.

Ο ορισμός που ακολουθεί έχει εισαχθεί από τους $Cellier$, $Fourdrinier$ και $Robert$[1986]

Ορισμός 1.5.1

Έστω $\lambda \in \mathbb{R}_+$. Μία απεικόνιση f, από τον $\mathbb{R}_+$ στον $\mathbb{R}_+$, θα λέγεται ελεγχόμενη σε ύψος λ αν η $t^\lambda f(t)$ είναι αύξουσα συνάρτηση.

1.5.2

Λαμβάνουμε το ακόλουθο αποτέλεσμα:

Πρόταση 1.5.6

Αν $rg(b) \geq 3$ και οι υποθέσεις H_1, H_2, H_4 ικανοποιούνται, μία ικανή συνθήκη ώστε ο εκτιμητής ϕ να κυριαρχεί τον εκτιμητή ϕ_0 είναι να υπάρχουν δύο πραγματικοί αριθμοί μ_1 και μ_2, *τέτοιοι ώστε, για κάθε* $u > 0$, *η* $h(\ .\ ,u)$ *να είναι ελεγχόμενη σε ύψος* μ_1, *για κάθε* $t > 0$, *η* $h(t,\ .\ )$ *να είναι ελεγχόμενη σε ύψος* μ_2 *και η* $tuh(t,u)$ *να είναι ομοιόμορφα φραγμένη από το:*

$$R_2 \ = \ 2\ \frac{tr(v_\Theta qc) \ - \ 2\ \mu_1\ \lambda_M}{\sigma_M} \quad \frac{n-k}{n-k+4\mu_2-2}$$

Παρατήρηση 1.5.8

Βασισμένο στον $Alam$[1973], που έχει εισαγάγει την έννοια του "$contr\widehat{ol}$" μέσω των συναρτήσεων δυνάμεως, το παραπάνω αποτέλεσμα είναι πολύ πλησίον σε αυτό των $Cellier$, $Fourdrinier$ και $Robert$[1989], όπου οι αλγεβρικές υποθέσεις επί των b και c είναι πιο περιοριστικές απ'οτι εδώ, αλλά όπου αντιθέτως ο συρρικνωτής δεν θεωρείται διαφορίσιμος.

Παρατήρηση 1.5.9

Έχουμε οτι $R_2 \leq R_0$. Επιπλέον, $R_2 = R_0$ στην περίπτωση που τα μ_1 και μ_2 είναι ίσα με 1, ελάχιστη τιμή για την οποία οι υποθέσεις της Πρότασης 1.5.6 είναι συμβατές. Για παράδειγμα, αν $\mu_2 < 1$, τότε

$$\forall t > 0\ , \quad \lim_{u\to+\infty} u\, h(t,u) \ = \ \lim_{u\to+\infty} u^{1-\mu_2}\ u^{\mu_2}\ h(t,u) = +\infty\ ,$$

το οποίο είναι άτοπο αφού η $uh(t,u)$ είναι φραγμένη.

Απόδειξη

Θα αποδείξουμε οτι ικανοποιούνται οι υποθέσεις της Πρότασης 1.4.3.

Η υπόθεση του φραγμένου της $tuh(t,u)$ αρκεί για να αποδειχθούν οι συνθήκες ομαλότητας του ϕ. Για να είναι η $\mathbb{E}_{\theta,\sigma}[\ th^2(t,1/s^2)\ ]$ πεπερασμένη, αρκεί να είναι πεπερασμένη η $\mathbb{E}_{\theta,\sigma}[\ s^4/t\ ]$, αφού: $\mathbb{E}_{\theta,\sigma}[\ th^2(t,1/s^2)\ ] = \mathbb{E}_{\theta,\sigma}[\ \frac{s^4}{t}\ t^2\ \frac{1}{s^4}\ h^2(t,1/s^2)\ ] \leq R_2^2\ \mathbb{E}_{\theta,\sigma}[\ s^4/t\ ]$. Ομοίως για το πεπερασμένο της $\mathbb{E}_{\theta,\sigma}[\ \frac{t}{s^2}h^2(t,1/s^2)\ ]$ αρκεί να είναι πεπερασμένη η $\mathbb{E}_{\theta,\sigma}[\ s^2/t\ ]$. Το πεπερασμένο αυτών των μέσων τιμών εξασφαλίζεται, όπως είδαμε στην §1.4.2.(B), από την ανεξαρτησία του t και του s^2 και το πεπερασμένο των $\mathbb{E}_{\theta,\sigma}[\ 1/t\ ]$, $\mathbb{E}_{\theta,\sigma}[\ s^2\ ]$, $\mathbb{E}_{\theta,\sigma}[\ s^4\ ]$. Επιπλέον, η υπόθεση του φραγμένου της $tuh(t,u)$, μας επιτρέπει να ολοκληρώσουμε κατά μέρη τα ολοκληρώματα στα οποία εφαρμόσαμε το Λήμμα του $Stein$ στην απόδειξη της Πρότασης 1.3.2 . Αντιστοίχως λοιπόν, για $1 \leq i \leq k$, έχουμε:

$$\begin{aligned}
&\int_{-\infty}^{+\infty} h(t,1/s^2)(d_iy_i+\beta_i)\frac{y_i-\theta_i}{\sigma^2}\ f_i(y_i)\,dy_i \\
= \ &\int_{-\infty}^{+\infty} \frac{\partial}{\partial y_i}\Big(h(t,1/s^2)(d_iy_i+\beta_i)\Big)\ f_i(y_i)\,dy_i \ - \ \Big[h(t,1/s^2)\ (d_iy_i+\beta_i)f_i(y_i)\Big]_{-\infty}^{+\infty}
\end{aligned}$$

Όμως,

$$\begin{aligned}
\Big|h(t,1/s^2)\ (d_iy_i+\beta_i)f_i(y_i)\Big| \ &= \ \Big|\frac{t}{s^2}\ h(t,1/s^2)\Big|\ \Big|\frac{(d_iy_i+\beta_i)s^2}{t}f_i(y_i)\Big| \\
&\leq \ R_2\ s^2\Big|\frac{(d_iy_i+\beta_i)s^2}{\sum_{j=1}^k b_jy_j^2}f_i(y_i)\Big| \ \longrightarrow\ 0 \ \ ,\text{όταν } y_i \to \pm\infty
\end{aligned}$$

Άρα, ακριβώς όπως στην απόδειξη της Πρότασης 1.3.2, προκύπτει τελικά οτι:

$$2\ \mathbb{E}_{\theta,\sigma}\Big[h_1'(t,1/s^2)(\overline{bv_\Theta qc})(\ \phi_0(y)\ )\Big] = A_\phi(\theta,\sigma) - tr(v_\Theta qc)\ \mathbb{E}_{\theta,\sigma}\Big[h(t,1/s^2)\Big]$$

και επειδή το δεύτερο μέλος είναι πεπερασμένο, χρησιμοποιώντας τη σχέση (1.5), προκύπτει οτι:

$$\mathbb{E}_{\theta,\sigma}\Big[\mid t\ h_1'(t,1/s^2)\mid\Big] < +\infty$$

Ομοίως, για $k+1 \leq i \leq n$, έχουμε:

$$\begin{aligned}
&\int_{-\infty}^{+\infty} \frac{1}{(n-k)s^2}\ h^2(t,1/s^2) y_i\ \frac{y_i}{\sigma^2}\ f_i(y_i)\, dy_i \\
=\ &\int_{-\infty}^{+\infty} \frac{\partial}{\partial y_i}\Big(h^2(t,1/s^2)\ \frac{y_i}{(n-k)s^2}\Big) f_i(y_i)\, dy_i\ -\ \Big[h^2(t,1/s^2)\ \frac{y_i}{(n-k)s^2}\ f_i(y_i)\Big]_{-\infty}^{+\infty}
\end{aligned}$$

Όμως,

$$\begin{aligned}
\Big|h^2(t,1/s^2)\ \frac{y_i}{(n-k)s^2}\ f_i(y_i)\Big| \ &=\ \Big|\frac{t^2}{s^4}\ h^2(t,1/s^2)\Big|\ \Big|\frac{s^2 y_i}{(n-k)t^2}\ f_i(y_i)\Big| \\
&\leq\ \frac{1}{(n-k)t^2}\ R_2^2 \Big|s^2 y_i f_i(y_i)\Big|\ \longrightarrow\ 0\ \ ,\text{όταν } y_i \to \pm\infty
\end{aligned}$$

Οπότε, όπως στην απόδειξη της Πρότασης 1.3.2, λαμβάνουμε:

$$\begin{aligned}
\frac{4}{n-k}\ \mathbb{E}_{\theta,\sigma}\Big[\frac{1}{s^4}\ h(t,1/s^2)\ h_2'(t,1/s^2)\ \bar{q}\Big(c(\phi_0(y))\Big)\Big] \ =\ &\mathbb{E}_{\theta,\sigma}\Big[\frac{n-k-2}{(n-k)s^2}\ h^2(t,1/s^2)\ \bar{q}\Big(c(\phi_0(y))\Big)\Big] \\
&-\ B_\phi(\theta,\sigma)
\end{aligned}$$

απ΄ όπου χρησιμοποιώντας τη σχέση (1.3) καταλήγουμε στο:

$$\mathbb{E}_{\theta,\sigma}\Big[\ \Big|\frac{t}{s^4}\ h(t,1/s^2)\ h_2'(t,1/s^2)\Big|\ \Big] < +\infty$$

Έχουμε, από τις υποθέσεις του "ελέγχου" για τη συνάρτηση συρρίκνωσης h, οτι $r(t,u)\ t^{\mu_1-1}\ u^{\mu_2-1}$ είναι αύξουσα συνάρτηση ξεχωριστά ως πρός t και ως πρός u. Συνεπώς:

$$r_1'(t,u)\ t^{\mu_1-1}\ u^{\mu_2-1}\ +\ (\ \mu_1\ -\ 1)\ r(t,u)\ t^{\mu_1-2}\ u^{\mu_2-1} \geq 0\ \Rightarrow\ r_1'(t,u) \geq -(\mu_1-1)\frac{r(t,u)}{t}$$

και ομοίως:

$$r_2'(t,u) \geq -(\mu_2-1)\frac{r(t,u)}{u}$$

Απ΄ όπου προκύπτει :

$$\begin{aligned}
\lambda_m r_1'^{+}(t,u) - \lambda_M r_1'^{-}(t,u) \ &=\ inf\Big\{\lambda_m r_1'(t,u)\ ,\ \lambda_M r_1'(t,u)\Big\} \\
&\geq\ -\lambda_M(\mu_1-1)\frac{r(t,u)}{t}
\end{aligned}$$

και

$$\begin{aligned}
\sigma_m r_2'^{+}(t,u) - \sigma_M r_2'^{-}(t,u) \ &=\ inf\Big\{\sigma_m r_2'(t,u)\ ,\ \sigma_M r_2'(t,u)\Big\} \\
&\geq\ -\sigma_M(\mu_2-1)\frac{r(t,u)}{u}
\end{aligned}$$

Λοιπόν, από τις παραπάνω σχέσεις και το γεγονός ότι, $r(t,u) \leq R_2$, παίρνουμε

$$\begin{aligned}
&\sigma_M \frac{n-k+2}{n-k} r(t,u)\Big(R_0 - r(t,u)\Big) \ + \ 4t\Big(\lambda_m r_1'^{+}(t,u) - \lambda_M r_1'^{-}(t,u)\Big) \\
&\qquad\qquad\qquad\qquad + 4\frac{r(t,u)u}{n-k}\Big(\sigma_m r_2'^{+}(t,u) - \sigma_M r_2'^{-}(t,u)\Big) \\
&\geq \ \sigma_M \frac{n-k+2}{n-k}\, r(t,u)\Big(R_0 - r(t,u)\Big) - 4\lambda_M(\mu_1 - 1) r(t,u) - 4\sigma_M \frac{\mu_2 - 1}{n-k}\, r^2(t,u) \\
&= \ r(t,u)\Big[\sigma_M \frac{n-k+2}{n-k} R_0 - \sigma_M \frac{n-k+2}{n-k}\, r(t,u) - 4\lambda_M(\mu_1 - 1) - 4\sigma_M \frac{\mu_2 - 1}{n-k}\, r(t,u)\Big] \\
&= \ r(t,u)\Big[2(\ tr(v_\Theta qc\ ) - 2\lambda_M\ ) - 4\lambda_M(\mu_1 - 1) - \ \frac{\sigma_M}{n-k}(n-k+4\mu_2 - 2) r(t,u)\Big] \\
&\geq \ 0
\end{aligned}$$

δηλαδή, ικανοποιείται και η υπόθεση (1.6) της προτάσεως 1.4.3. Επομένως, ο ϕ κυριαρχεί τον ϕ_0. $\square$

1.6 ΠΑΡΑΡΤΗΜΑ

Έστω V ένας διανυσματικός χώρος, διάστασης n, επί του $\mathbb{R}$ και a μία διγραμμική μορφή επί του V. Σημειώνουμε με ${}^s a \ = \ \frac{1}{2}\ (a \ + \ {}^t a)$ την διγραμμική συμμετρική μορφή και με $\bar{a}$ την τετραγωνική μορφή που συνδέεται με την a.

Αν a είναι μία διγραμμική συμμετρική μορφή επί του V, μπορούμε να την θεωρήσουμε σαν έναν ομομορφισμό του V μέσα στον V^*(δυϊκός του V). Σημειώνουμε με a^- έναν γενικευμένο αντίστροφο του a, δηλαδή έναν ομομορφισμό του V^* μέσα στον V, που ικανοποιεί τη συνθήκη: $a\, a^-\, a = a$ (δηλ. $a^-|_{Im a} = (a|_{E'})^{-1}$, όπου E' ο συμπληρωματικός υπόχωρος του $Ker\, a$). Έχουμε τότε $Im\, a^- \oplus Ker\, a = V$.

Υπενθυμίζουμε στο παρακάτω Λήμμα κάποιες χρήσιμες αλγεβρικές ιδιότητες.

Λήμμα 1.6.1

Έστω a και b δύο διγραμμικές μορφές επί ενός διανυσματικού χώρου V διαστάσεως n. Η b είναι συμμετρική και θετική. Υποθέτουμε οτι: $Ker\, b \subset Ker\, a$ και $Ker\, b \subset Ker\, {}^t a$. Τότε:

(i) $Im\, a \subset Im\, b, \ \ Im\, {}^t a \subset Im\, b, \ \ Im\, {}^s a \subset Im\, b$

(ii) Οι περιορισμοί στην $Im\, b$ των ab^-, ${}^t ab^-$ και ${}^s ab^-$ δεν εξαρτώνται από την επιλογή του γενικευμένου αντίστροφου b^- του b.

(iii) Τα φάσματα των ab^- και ${}^t ab^-$ συμπίπτουν και δεν εξαρτώνται απο την επιλογή του b^-. Το κοινό φάσμα διαχωρίζεται στο Λ, φάσμα του περιορισμού στην $Im\, b$ και σε $n - rg(b)$ μηδενικές ιδιοτιμές.

(iv) $$\inf_{x \notin Ker\, b} \frac{\bar{a}(x)}{\bar{b}(x)} \ = \ inf\, \Lambda \quad \text{και} \quad \sup_{x \notin Ker\, b} \frac{\bar{a}(x)}{\bar{b}(x)} \ = \ sup\, \Lambda$$

Απόδειξη

(i) Επειδή $Ker\, b \subset Ker\, a$, έχουμε οτι $dim(Ker\, b) \leq dim(Ker\, a)$. Άρα,

$$dim(Im\, a) \;=\; n \;-\; dim(Ker\, a) \;\leq\; n \;-\; dim(Ker\, b) \;=\; dim(Im\, b)$$

απ΄ όπου προκύπτει οτι: $Im\, a \subset Im\, b$.
Όμοια για τις άλλες περιπτώσεις.

(ii) Έστω b_1^-, b_2^- δύο γενικευμένοι αντίστροφοι του b, τότε, για κάθε $y \in Im\, b$,

$$b_1^-(y) - b_2^-(y) \in Ker\, b (\subset Ker\, a)$$

Άρα, $a\Big(b_1^-(y) - b_2^-(y)\Big) = ab_1^-(y) - ab_2^-(y) = 0 \;\Rightarrow\; ab_1^-(y) = ab_2^-(y),\ \forall y \in Im\, b$

(iii) Προφανές, λόγω του οτι τα φάσματα των a και ${}^t a$ ταυτίζονται.

(iv) Δεδομένου οτι η b είναι μία διγραμμική, συμμετρική μορφή, υπάρχει ένα ζεύγος δυϊκών βάσεων $\mathcal{E} = (e_1, \ldots, e_n)$ και $\mathcal{E}^* = (e_1^*, \ldots, e_n^*)$, $b-$ορθοκανονικών όπου τα $e_{m+1}, \ldots, e_n$ να αποτελούν βάση για τον $Ker\, b$ ($m = rg(b)$). Ο πίνακας της b ως προς το ζεύγος αυτό των βάσεων είναι:

$$\begin{pmatrix} I_m & 0 \\ 0 & 0 \end{pmatrix}$$

Ο πίνακας ενός γενικευμένου αντιστρόφου b^- του b θα είναι :

$$\begin{pmatrix} I_m & \alpha \\ \beta & \gamma \end{pmatrix}$$

όπου α, β, γ πίνακες τύπου $(m, n-m)$, $(n-m, m)$ και $(n-m, n-m)$ αντιστοίχως.

Έστω (a_{ij}) ο πίνακας του a αναφορικά με αυτές τις βάσεις. Επειδή $Ker\, b \subset Ker\, a$ και $Ker\, b \subset Ker {}^t a$, $a_{ij} = 0$ για $i > m$ ή $j > m$, δηλαδή:

$$(a_{ij}) = \begin{pmatrix} A_m & 0 \\ 0 & 0 \end{pmatrix} \quad \Big(= \ {}_{\mathcal{E}^*}(a)_{\mathcal{E}} \ \Big)$$

Άρα,

$${}_{\mathcal{E}^*}(\ ab^-)_{\mathcal{E}^*} = \begin{pmatrix} A_m & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} I_m & \alpha \\ \beta & \gamma \end{pmatrix} = \begin{pmatrix} A_m & A_m\alpha \\ 0 & 0 \end{pmatrix}$$

και

$${}_{\mathcal{E}^*}({}^t ab^-)_{\mathcal{E}^*} = \begin{pmatrix} {}^t A_m & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} I_m & \alpha \\ \beta & \gamma \end{pmatrix} = \begin{pmatrix} {}^t A_m & {}^t A_m\alpha \\ 0 & 0 \end{pmatrix}$$

Παρατηρούμε οτι: $tr(ab^-) = tr\, A_m = tr\, {}^t A_m = tr({}^t ab^-) = tr({}^s ab^-)$, $\forall b^-$ γενικευμένο αντίστροφο του b.

Εκλέγοντας την βάση $\mathcal{E}$ να είναι και ${}^s a$-ορθογώνια, ο πίνακας του ${}^s ab^-$ περιορισμένου στην $Im\, b$ είναι τότε διαγώνιος και δέχεται για φάσμα το $\Lambda = \{\upsilon_1, \ldots, \upsilon_m\}$.

Έστω $x \in V$ και $x_1, \ldots, x_n$ οι συντεταγμένες του ως πρός τη βάση $\mathcal{E}$. Τότε:

$$\bar{a}(x) = \sum_{i=1}^{m} \upsilon_i x_i^2 \quad \text{και} \quad \bar{b}(x) = \sum_{i=1}^{m} x_i^2$$

Επομένως,

$$\min_{1 \leq i \leq m} \upsilon_i \leq \frac{\bar{a}(x)}{\bar{b}(x)} = \frac{\sum_{i=1}^{m} \upsilon_i x_i^2}{\sum_{i=1}^{m} x_i^2} \leq \max_{1 \leq i \leq m} \upsilon_i$$

□

Παρατήρηση 1.6.10

Το παραπάνω Λήμμα εφαρμόζεται ιδιαιτέρως για $a = bd$, όπου d είναι ένας ενδομορφισμός του V τέτοιος ώστε $Ker\, b \subset Ker\, d$. Επαληθεύεται στοιχειωδώς οτι το φάσμα του bdb^{-} συμπίπτει με το φάσμα του d.

Πράγματι, επειδή $Ker\, b \subset Ker\, d$, ο πίνακας του d ως πρός τη βάση $\mathcal{E}$ έχει τις $n-m$ τελευταίες στήλες μηδέν, δηλαδή,

$$ {}_{\mathcal{E}}(d)_{\mathcal{E}} = \begin{pmatrix} C_m & 0 \\ * & 0 \end{pmatrix} $$

Ο πίνακας του bd ως πρός το ζεύγος των παραπάνω δυϊκών βάσεων $\mathcal{E}$ και $\mathcal{E}^*$ συμπεραίνεται από τον ${}_{\mathcal{E}}(d)_{\mathcal{E}}$ αντικαθιστώντας τις $n-m$ τελευταίες γραμμές με το μηδέν, αφού:

$$ {}_{\mathcal{E}^*}(bd)_{\mathcal{E}} = {}_{\mathcal{E}^*}(b)_{\mathcal{E}}\ {}_{\mathcal{E}}(d)_{\mathcal{E}} = \begin{pmatrix} I_m & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} C_m & 0 \\ * & 0 \end{pmatrix} = \begin{pmatrix} C_m & 0 \\ 0 & 0 \end{pmatrix} $$

Το ίδιο γίνεται για εκείνον του bdb^{-}, αφού:

$$ {}_{\mathcal{E}^*}(bdb^{-})_{\mathcal{E}^*} = {}_{\mathcal{E}^*}(bd)_{\mathcal{E}}\ {}_{\mathcal{E}}(b^{-})_{\mathcal{E}^*} = \begin{pmatrix} C_m & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} I_m & \alpha \\ \beta & \gamma \end{pmatrix} = \begin{pmatrix} C_m & C_m\alpha \\ 0 & 0 \end{pmatrix} . $$

Το παρακάτω λήμμα οφείλεται στον $Stein$[1981]:

Λήμμα 1.6.2 ***(του* Stein*)***

Έστω Y μία τυχαία μεταβλητή που ακολουθεί έναν κανονικό νόμο μέσου θ και διασποράς σ^2 και g μία συνάρτηση από τον $\mathbb{R}$ στον $\mathbb{R}$, διαφορίσιμη και τέτοια ώστε $\mathbb{E}|\,g'(Y)|$ να είναι πεπερασμένη. Τότε:

$$ \mathbb{E}\Big[\ \frac{Y-\theta}{\sigma^2}\ g(Y)\ \Big] = \mathbb{E}\Big[\ g'(Y)\ \Big] $$

Απόδειξη

Συμβολίζοντας με $f(y)$ την πυκνότητα της κανονικής κατανομής $N(\theta,\ \sigma^2)$, η παράγωγος της είναι:

$$ f'(y) = -\frac{y-\theta}{\sigma^2}\ f(y) $$

Λοιπόν,

$$ \begin{aligned} \mathbb{E}\Big[\ g'(Y)\ \Big] &= \int_{-\infty}^{+\infty} g'(y)\ f(y)\ dy \\ &= \int_0^{+\infty} g'(y)\Big(\int_y^{+\infty} \frac{z-\theta}{\sigma^2}\ f(z)\ dz\Big)dy - \int_{-\infty}^{0} g'(y)\Big(\int_{-\infty}^{y} \frac{z-\theta}{\sigma^2}\ f(z)\ dz\Big)dy \\ &= \int_0^{+\infty} \frac{z-\theta}{\sigma^2} f(z)\Big(\int_0^{z} g'(y)\ dy\Big)dz - \int_{-\infty}^{0} \frac{z-\theta}{\sigma^2} f(z)\Big(\int_z^{0} g'(y)\ dy\Big)dz \\ &= \int_0^{+\infty} \frac{z-\theta}{\sigma^2} f(z)\Big(g(z)-g(0)\Big)dz + \int_{-\infty}^{0} \frac{z-\theta}{\sigma^2} f(z)\Big(g(z)-g(0)\Big)dz \end{aligned} $$

$$
\begin{aligned}
&= \int_{-\infty}^{+\infty} \frac{z-\theta}{\sigma^2} f(z)\Big(g(z)-g(0)\Big)dz \\
&= \int_{-\infty}^{+\infty} \frac{z-\theta}{\sigma^2} f(z)\; g(z)\; dz \;-g(0) \underbrace{\int_{-\infty}^{+\infty} f'(z)dz}_{=0} \\
&= \int_{-\infty}^{+\infty} \frac{z-\theta}{\sigma^2} g(z)\; f(z)\; dz \\
&= \mathbb{E}\Big[\; \frac{Y-\theta}{\sigma^2}\; g(Y)\; \Big]
\end{aligned}
$$

Η τρίτη ισότητα παραπάνω προέκυψε εφαρμόζοντας το Θεώρημα του $Fubini$ και στους δύο όρους, την χρήση του οποίου μας εξασφαλίζει η υπόθεση του Λήμματος: $\mathbb{E}|\,g'(Y)| < +\infty$, απ΄ όπου προκύπτει:

$$
\int_{0}^{+\infty} |\,g'(y)| f(y)\; dy \leq \mathbb{E}|\,g'(Y)| < +\infty \quad \text{και} \quad \int_{-\infty}^{0} |\,g'(y)| f(y)\; dy \leq \mathbb{E}|\,g'(Y)| < +\infty
$$

$\square$

Κεφάλαιο 2

ΜΙΑ ΑΝΑΓΚΑΙΑ ΣΥΝΘΗΚΗ ΠΑΡΑΔΟΧΗΣ ΚΑΙ ΤΑ ΑΠΟΤΕΛΕΣΜΑΤΑ ΤΗΣ ΕΠΙ ΤΩΝ ΕΚΤΙΜΗΤΩΝ ΜΕ ΣΥΡΡΙΚΝΩΤΗ

2.1

Θεωρώντας εδώ τις εκθετικές οικογένειες πυκνοτήτων, θα εκτιμήσουμε ενδιαφέρουσες παραμέτρους που δεν είναι οι φυσικές παράμετροι. Ακολουθώντας το [33], αποδεικνύουμε, υπό μια υπόθεση συμπαγότητας, οτι οι παραδεκτοί εκτιμητές αυτών των παραμέτρων είναι όρια εκτιμητών του *Bayes* και μπορούν να εκφρασθούν υπό μια μορφή ιδιαίτερα συναρτησιακή. Μια ενδιαφέρουσα ειδική περίπτωση στο πλαίσιο αυτής της μελέτης οδηγεί στην εκτίμηση του μέσου μιας πολυδιάστατης κανονικής κατανομής όταν η διασπορά είναι γνωστή εκτός από έναν πολλαπλασιαστικό παράγοντα. Συμπεραίνουμε από το γενικό αποτέλεσμα μια αναγκαία συνθήκη για την παραδοχή των εκτιμητών με συρρικνωτή και εργαζόμενοι σε μία υποκλάση συμπεραίνουμε, από το Κεφ.1, ικανές συνθήκες ομοιόμορφης κυριαρχίας επί του ε.ε.τ.

2.2 ΤΟ ΜΟΝΤΕΛΟ

Οι έννοιες που θα χρησιμοποιηθούν παρακάτω είναι αυτές της παραγράφου 1.2.1.

2.2.1 Εκθετικές Οικογένειες

Θεωρούμε εδώ, τις εκθετικές οικογένειες $(k+1)-$διάστασης του παρακάτω τύπου: Έστω Θ ένας ευκλείδειος χώρος διάστασης k επί του $\mathbb{R}$ και έστω Δ ένα συμπαγές διάστημα του $\mathbb{R}_+^*$. Παρατηρούμε το τυχαίο ζεύγος (u,z) του $\Theta\times\mathbb{R}$ με νόμο που δέχεται πυκνότητα αναφορικά με ένα μέτρο λ, σ-πεπερασμένο, την:

$$p(u,z,\theta,\delta) = exp\Big\{\theta.u + \delta z - K(\theta,\delta)\Big\}$$

Σημειώνουμε με H την κλειστή, κυρτή θήκη του φορέα του λ.

Η παράμετρος (θ,δ) είναι άγνωστη και ενδιαφερόμαστε για την εκτίμηση του θ/δ. Οι εκτιμητές του θ/δ, δηλ. οι μετρήσιμες απεικονίσεις από τον $\Theta\times\mathbb{R}$ στον Θ, συγκρίνονται μέσω μιας συνάρτησης απωλείας που συνδέεται με μια θετικά ορισμένη, συμμετρική διγραμμική μορφή q επί του Θ, έστω

$$L\Big(d,(\theta,\delta)\Big) = \delta\,\bar{q}(d-\frac{\theta}{\delta})$$

2.2.2 Ειδική Περίπτωση των Εκτιμητών του Bayes

Έστω ρ ένα πεπερασμένο μέτρο επί του $\Theta \times \Delta$, με συμπαγή φορέα. Σημειώνουμε μ το μέτρο επί του $\Theta \times \Delta$, που δέχεται πυκνότητα την $exp\{-K(\theta,\delta)\}$ αναφορικά με το ρ και f_μ ο μετασχηματισμός *Laplace* του μ, για το βαθμωτό γινόμενο ορισμένο επί του $\Theta \times \mathbb{R}$ από τη σχέση: $<(\theta,\delta),(u,z)>= \theta.u + \delta z$, δηλαδή

$$f_\mu(u,z) = \int_{\Theta\times\Delta} exp\Big\{ <(\theta,\delta),(u,z)> \Big\} \mu(d\theta,d\delta)$$

Προκύπτει τότε από τη μορφή της συνάρτησης απωλείας οτι ο εκτιμητής:

$$\phi_\rho(u,z) = \frac{\int \theta\, exp(\theta.u + \delta z)\mu(d\theta,d\delta)}{\int \delta\, exp(\theta.u + \delta z)\mu(d\theta,d\delta)} = \frac{\frac{\partial f_\mu}{\partial u}(u,z)}{\frac{\partial f_\mu}{\partial z}(u,z)}$$

είναι ένας εκτιμητής του *Bayes* για το θ/δ αναφορικά με το εκ των προτέρων μέτρο ρ. (η $\frac{\partial f_\mu}{\partial u}(u,z)$ ορίζεται ως ένα στοιχείο του Θ σύμφωνα με την ευκλείδεια δομή αυτού του χώρου)

Πράγματι, επειδή το ρ είναι πεπερασμένο και με συμπαγή φορέα, η f_μ είναι ορισμένη σε ολόκληρο τον $\Theta \times \mathbb{R}$, διαφορίσιμη και επαληθεύει:

$$\forall (u,z) \in \Theta \times \mathbb{R}\,, \qquad \frac{\partial f_\mu}{\partial u}(u,z) = \int_{\Theta\times\Delta} \theta\, exp\{\theta.u + \delta\, z\}\mu(d\theta,d\delta)$$
$$\text{και} \qquad \frac{\partial f_\mu}{\partial z}(u,z) = \int_{\Theta\times\Delta} \delta\, exp\{\theta.u + \delta\, z\}\mu(d\theta,d\delta)$$

Σημειώνοντας με $R_\phi(\theta,\delta)$ τον (κλασσικό) κίνδυνο ενός εκτιμητή ϕ του θ/δ,

$$R_\phi(\theta,\delta) = \int_{\Theta\times\mathbb{R}} L\Big(\phi(u,z) - (\theta,\delta)\Big) p(u,z,\theta,\delta)\lambda(du,dz)\,,$$

ο *bayesien* κίνδυνος του ϕ αναφορικά με το εκ των προτέρων μέτρο ρ,

$$r_\phi(\rho) = \int_{\Theta\times\Delta} R_\phi(\theta,\delta)\rho(d\theta,d\delta)$$

σύμφωνα με το Θεώρημα του *Fubini*, γίνεται:

$$r_\phi(\rho) = \int_{\Theta\times\mathbb{R}} \left(\int_{\Theta\times\Delta} \delta\, \bar{q}(\phi(u,z) - \theta/\delta) p(u,z,\theta,\delta)\rho(d\theta,d\delta) \right) \lambda(du,dz)$$

Γνωρίζοντας λοιπόν οτι ο εκτιμητής του *Bayes* του θ/δ, ελαχιστοποιεί τον *bayesien* κίνδυνο, αυτός θα προκύψει συνδέοντας με το (u,z) το σημείο $\phi_\rho(u,z)$ του Θ στο οποίο η συνάρτηση

$$\Phi_{(u,z)}(d) = \int_{\Theta\times\Delta} \delta\, \bar{q}(d - \theta/\delta) p(u,z,\theta,\delta)\rho(d\theta,d\delta)$$

λαμβάνει την ελάχιστη τιμή της (υπό την προϋπόθεση οτι ένα τέτοιο σημείο υπάρχει για κάθε (u,z)).

Για τον σκοπό αυτό έχουμε:

$$\begin{aligned}
\Phi_{(u,z)}(d) &= \int_{\Theta\times\Delta} \delta\,\bar{q}(d-\theta/\delta)exp\{\theta . u+\delta\, z\}\mu(d\theta,d\delta)\\
&= \bar{q}(d)\int_{\Theta\times\Delta}\delta\; exp\{\theta . u+\delta\, z\}\mu(d\theta,d\delta)+\int_{\Theta\times\Delta}\delta^{-1}\bar{q}(\theta)exp\{\theta . u+\delta\, z\}\mu(d\theta,d\delta)\\
&\quad -2q\Big(d,\int_{\Theta\times\Delta}\theta\; exp\{\theta . u+\delta\, z\}\mu(d\theta,d\delta)\Big)
\end{aligned}$$

και

$$\Phi'_{(u,z)}(d) = 2q\,d\int_{\Theta\times\Delta}\delta\; exp\{\theta . u+\delta\, z\}\mu(d\theta,d\delta)-2q\int_{\Theta\times\Delta}\theta\; exp\{\theta . u+\delta\, z\}\mu(d\theta,d\delta)$$

Επειδή η q είναι θετικά ορισμένη, η εξίσωση $\Phi'_{(u,z)}(d)=0$ οδηγεί στην λύση:

$$\widehat{d}=\frac{\int_{\Theta\times\Delta}\theta\; exp\{\theta . u+\delta\, z\}\mu(d\theta,d\delta)}{\int_{\Theta\times\Delta}\delta\; exp\{\theta . u+\delta\, z\}\mu(d\theta,d\delta)}$$

Παρατηρώντας οτι, η

$$\Phi''_{(u,z)}(d)=\Big(2\int_{\Theta\times\Delta}\delta\; exp\{\theta . u+\delta\, z\}\mu(d\theta,d\delta)\Big)q$$

είναι θετικά ορισμένη, ο $\phi_\rho(u,z)=\widehat{d}$ είναι ο ζητούμενος εκτιμητής.

2.3 ΠΑΡΑΔΟΧΗ

Το αποτέλεσμα της παραδοχής που θα αναφέρουμε παρακάτω επεκτείνει, με την εκτίμηση του θ/δ, τη μορφή της αναγκαίας συνθήκης για την παραδοχή των εκτιμητών των φυσικών παραμέτρων (θ,δ), που αποδείχθηκε από τον *Brown*[1986].

Για κάθε μέτρο μ επί του χώρου των παραμέτρων $\Theta\times\Delta$, το Ψ_μ θα σημειώνει τον λογάριθμο του μετασχηματισμού *Laplace* του μέτρου αυτού.

Λήμμα

Αν (μ_n) είναι μια ακολουθία μέτρων με πεπερασμένο φορέα, τέτοια ώστε, για σχεδόν κάθε (u,z), η οικογένεια $(\|grad\ \Psi_{\mu_n}(u,z)\|)$ να είναι φραγμένη, τότε υπάρχει ένα μέτρο μ και μια υπακολουθία (μ_{n_k}) τέτοια ώστε για σχεδόν κάθε (u,z) του H^o, να έχουμε:

$$\lim_{k\to+\infty}\Psi_{\mu_{n_k}}(u,z)=\Psi_\mu(u,z)\ \text{ και }\ \lim_{k\to+\infty} grad\,\Psi_{\mu_{n_k}}(u,z)=grad\,\Psi_\mu(u,z)$$

Το λήμμα αυτό περιέχεται στην απόδειξη του Θεωρήματος 4.14 του *Brown*[1986] η οποία βασίζεται στο Θεώρημα 2.17 του ίδιου άρθρου (συνέχεια του μετασχηματισμού *Laplace* ενός μέτρου). Το Λήμμα αυτό είναι σημαντικό για την απόδειξη της παρακάτω Πρότασης:

Πρόταση 2.3.1

Αν ο ϕ είναι ένας παραδεκτός εκτιμητής του θ/δ, όπου το δ ανήκει σε ένα συμπαγές διάστημα Δ του $\mathbb{R}_+^*$, υπάρχει ένα μέτρο μ_0 ορισμένο επί του $\Theta \times \Delta$, με μετασχηματισμό Laplace f_{μ_0}, τέτοιο ώστε για σχεδόν κάθε $(u, z) \in H^o$,

$$\phi(u, z) = \frac{\int \theta \, exp\{\theta.u + \delta \, z\}\mu_0(d\theta, d\delta)}{\int \delta \, exp\{\theta.u + \delta \, z\}\mu_0(d\theta, d\delta)} = \frac{\frac{\partial f_{\mu_0}}{\partial u}(u, z)}{\frac{\partial f_{\mu_0}}{\partial z}(u, z)}$$

Απόδειξη

Αφού ο ϕ είναι παραδεκτός προκύπτει, από το Θεώρημα 4.14 του *Brown*[1986] (του οποίου η εφαρμογή για την απώλεια $L(d, (\theta, \delta))$ είναι άμεση), οτι υπάρχει μια ακολουθία εκ των προτέρων μέτρων (ρ_n) με πεπερασμένο φορέα επί του $\Theta \times \Delta$ τέτοια ώστε η ακολουθία των εκτιμητών του *Bayes* ϕ_{ρ_n} να συγκλίνει λ-σχεδόν παντού προς τον ϕ. Με τις έννοιες της §2.2.2, έχουμε:

$$\phi_{\rho_n}(u, z) = \frac{\int \theta \, exp\{\theta.u + \delta \, z\}\mu_n(d\theta, d\delta)}{\int \delta \, exp\{\theta.u + \delta \, z\}\mu_n(d\theta, d\delta)}$$

με $d\mu_n = exp\{-K(\theta, \delta)\}d\rho_n$.

Έστω (u, z) τέτοιο ώστε

$$\lim_{n \to +\infty} \phi_{\rho_n}(u, z) = \phi(u, z)$$

και έστω c ένας πραγματικός αριθμός, τέτοιος ώστε $\|\phi_{\rho_n}(u, z)\| \leq c$. Σημειώνουμε με A το *supremum* του Δ. Έχουμε,

$$\begin{aligned}
&\left\| \int (\theta, \delta) exp\{\theta.u + \delta \, z\}\mu_n(d\theta, d\delta) \right\| \\
&\leq \left\| \int \theta \, exp\{\theta.u + \delta \, z\}\mu_n(d\theta, d\delta) \right\| + \left| \int \delta \, exp\{\theta.u + \delta \, z\}\mu_n(d\theta, d\delta) \right|
\end{aligned}$$

απ΄ όπου προκύπτει

$$\begin{aligned}
&\frac{\| \int (\theta, \delta) exp\{\theta.u + \delta \, z\}\mu_n(d\theta, d\delta) \|}{\int exp\{\theta.u + \delta \, z\}\mu_n(d\theta, d\delta)} \\
&\leq \frac{\| \int \theta \, exp\{\theta.u + \delta \, z\}\mu_n(d\theta, d\delta \|}{\int exp\{\theta.u + \delta \, z\}\mu_n(d\theta, d\delta)} + \frac{| \int \delta \, exp\{\theta.u + \delta \, z\}\mu_n(d\theta, d\delta) |}{\int exp\{\theta.u + \delta \, z\}\mu_n(d\theta, d\delta)} \\
&= \frac{| \int \delta \, exp\{\theta.u + \delta \, z\}\mu_n(d\theta, d\delta) |}{\int exp\{\theta.u + \delta \, z\}\mu_n(d\theta, d\delta)} \left[\frac{\| \int \theta \, exp\{\theta.u + \delta \, z\}\mu_n(d\theta, d\delta) \|}{| \int \delta \, exp\{\theta.u + \delta \, z\}\mu_n(d\theta, d\delta) |} + 1 \right] \\
&\leq A(c + 1)
\end{aligned}$$

Δηλαδή, $\|grad \, \Psi_{\mu_n}(u, z)\| \leq A(c + 1)$.

Άρα, οι υποθέσεις του παραπάνω Λήμματος ικανοποιούνται και συνεπώς έχουμε οτι υπάρχει ένα μέτρο μ_0 και μια υπακολουθία (μ_{n_k}) τέτοια ώστε:

$$\lim_{k \to +\infty} grad \, \Psi_{\mu_{n_k}}(u, z) = grad \, \Psi_{\mu_0}(u, z) \text{ , για σχεδόν κάθε } (u, z) \in H^o$$

δηλαδή,

$$\lim_{k \to +\infty} \frac{\int (\theta, \delta) exp\{\theta.u + \delta \, z\}\mu_{n_k}(d\theta, d\delta)}{\int exp\{\theta.u + \delta \, z\}\mu_{n_k}(d\theta, d\delta)} = \frac{\int (\theta, \delta) exp\{\theta.u + \delta \, z\}\mu_0(d\theta, d\delta)}{\int exp\{\theta.u + \delta \, z\}\mu_0(d\theta, d\delta)}$$

για σχεδόν κάθε $(u,z) \in H^o$.

Επομένως, για σχεδόν κάθε $(u,z) \in H^o$,

$$\lim_{k\to+\infty} \frac{\int \theta\ exp\{\theta.u + \delta\, z\}\mu_{n_k}(d\theta, d\delta)}{\int exp\{\theta.u + \delta\, z\}\mu_{n_k}(d\theta, d\delta)} = \frac{\int \theta\ exp\{\theta.u + \delta\, z\}\mu_0(d\theta, d\delta)}{\int exp\{\theta.u + \delta\, z\}\mu_0(d\theta, d\delta)} \tag{2.1}$$

και

$$\lim_{k\to+\infty} \frac{\int \delta\ exp\{\theta.u + \delta\, z\}\mu_{n_k}(d\theta, d\delta)}{\int exp\{\theta.u + \delta\, z\}\mu_{n_k}(d\theta, d\delta)} = \frac{\int \delta\ exp\{\theta.u + \delta\, z\}\mu_0(d\theta, d\delta)}{\int exp\{\theta.u + \delta\, z\}\mu_0(d\theta, d\delta)} \tag{2.2}$$

Επειδή το Δ είναι συμπαγές διάστημα του $\mathbb{R}_+^*$ έχουμε, για κάθε k, και σημειώνοντας με B το $infimum$ του Δ, οτι:

$$B \le \frac{\int \delta\ exp\{\theta.u + \delta\, z\}\mu_{n_k}(d\theta, d\delta)}{\int exp\{\theta.u + \delta\, z\}\mu_{n_k}(d\theta, d\delta)} \le A$$

δηλαδή,

$$\frac{\int \delta\ exp\{\theta.u + \delta\, z\}\mu_{n_k}(d\theta, d\delta)}{\int exp\{\theta.u + \delta\, z\}\mu_{n_k}(d\theta, d\delta)} \in \Delta$$

Λόγω της (2.2) και της συμπαγότητας του Δ, λαμβάνουμε

$$\frac{\int \delta\ exp\{\theta.u + \delta\, z\}\mu_0(d\theta, d\delta)}{\int exp\{\theta.u + \delta\, z\}\mu_0(d\theta, d\delta)} \in \Delta \tag{2.3}$$

Συμπεραίνουμε λοιπόν από τις (2.1), (2.2) και (2.3), οτι για σχεδόν κάθε $(u,z) \in H^o$,

$$\lim_{k\to+\infty} \frac{\int \theta\ exp\{\theta.u + \delta\, z\}\mu_{n_k}(d\theta, d\delta)}{\int \delta\ exp\{\theta.u + \delta\, z\}\mu_{n_k}(d\theta, d\delta)} = \frac{\int \theta\ exp\{\theta.u + \delta\, z\}\mu_0(d\theta, d\delta)}{\int \delta\ exp\{\theta.u + \delta\, z\}\mu_0(d\theta, d\delta)}$$

Δηλαδή, για σχεδόν κάθε $(u,z) \in H^o$, έχουμε:

$$\phi(u,z) = \frac{\int \theta\ exp\{\theta.u + \delta\, z\}\mu_0(d\theta, d\delta)}{\int \delta\ exp\{\theta.u + \delta\, z\}\mu_0(d\theta, d\delta)} = \frac{\frac{\partial f_{\mu_0}}{\partial u}(u,z)}{\frac{\partial f_{\mu_0}}{\partial z}(u,z)}$$

$\square$

2.4 ΕΦΑΡΜΟΓΗ ΣΤΟΝ ΚΑΝΟΝΙΚΟ ΝΟΜΟ

2.4.1

Όπως αναφέραμε και στην εισαγωγή, η κανονική κατανομή αποτελεί μια ειδική περίπτωση του μοντέλου που εισαγάγαμε στην παράγραφο 2.2. Έστω y ένα τυχαίο διάνυσμα με τιμές σε έναν διανυσματικό χώρο E, διάστασης n, επί του $\mathbb{R}$ που ακολουθεί μια κανονική κατανομή $N(m, \sigma^2 v)$, όπου $m \in \Theta$, Θ ένας διανυσματικός υπόχωρος του E, διάστασης k $(3 \le k \le n)$ και v μια θετικά ορισμένη, συμμετρική, διγραμμική μορφή στον E^*(δϋικός του E). Οι παράμετροι m και $\sigma^2(\sigma^2 > 0)$ είναι άγνωστοι.

Έχουμε οτι η v^{-1} ορίζει επί του E και λοιπόν επί του Θ, μια ευκλείδεια δομή και σημειώνουμε με ν το μέτρο του $Lebesgue$ που συνδέεται με αυτήν.

Αν ϕ_0 είναι η v^{-1}-ορθογώνια προβολή επί του Θ(δηλαδή, ο εκτιμητής ελαχίστων τετραγώνων του m), ο νόμος του y δέχεται, αναφορικά με το ν, πυκνότητα την:

$$\begin{aligned} p_{\theta,\sigma}(y) &= \frac{1}{(2\pi\sigma^2)^{n/2}}\, exp\Big\{-\frac{1}{2\sigma^2}\|y-m\|^2\Big\} \\ &= \frac{1}{(2\pi\sigma^2)^{n/2}}\, exp\Big\{-\frac{1}{2\sigma^2}\|y-\phi_0(y)+\phi_0(y)-m\|^2\Big\} \\ &= \frac{1}{(2\pi\sigma^2)^{n/2}}\, exp\Big\{-\frac{1}{2\sigma^2}\Big(\|y-\phi_0(y)\|^2+\|\phi_0(y)-m\|^2\Big)\Big\} \end{aligned}$$

Όμως,

$$\begin{aligned} &\|y-\phi_0(y)\|^2+\|\phi_0(y)-m\|^2 \\ &= \|y\|^2 - 2v^{-1}(y,\phi_0(y)) + 2\|\phi_0(y)\|^2 - 2v^{-1}(\phi_0(y),m) + \|m\|^2 \\ &= \|y\|^2 - 2\Big(v^{-1}(y,\phi_0(y)) - v^{-1}(\phi_0(y),\phi_0(y))\Big) - 2v^{-1}(\phi_0(y),m) + \|m\|^2 \\ &= \|y\|^2 - 2\underbrace{v^{-1}(y-\phi_0(y),\phi_0(y))}_{=\,0} - 2v^{-1}(\phi_0(y),m) + \|m\|^2 \\ &= -2v^{-1}(\phi_0(y),m) + \|y\|^2 + \|m\|^2 \end{aligned}$$

Συνεπώς, η $p_{\theta,\sigma}$ γράφεται,

$$p_{\theta,\sigma}(y) = exp\Big\{\frac{1}{\sigma^2}\ v^{-1}(\phi_0(y),m) - \frac{1}{2\sigma^2}\|y\|^2 - \frac{1}{2\sigma^2}\|m\|^2 - \frac{n}{2}\ \ln(2\pi\sigma^2)\Big\}$$

ή

$$p_{\theta,\sigma}(y) = exp\Big\{\frac{1}{\sigma^2}\ m.\phi_0(y) - \frac{1}{2\sigma^2}\|y\|^2 - \frac{1}{2\sigma^2}\|m\|^2 - \frac{n}{2}\ \ln(2\pi\sigma^2)\Big\}$$

αυτό το οποίο μας επιτρέπει να υποδείξουμε τον σύνδεσμο με τις έννοιες της παραγράφου 2.2:

$$u = \phi_0(y)\ ,\quad z = -\frac{1}{2}\|y\|^2 = t(y)\ ,\quad H = \Theta \times \mathbb{R}$$

λ είναι η εικόνα του ν υπό την απεικόνιση η οποία στο y αντιστοιχεί το (u,z). Η φυσική παράμετρος είναι το $(\theta,\delta) = (m/\sigma^2, 1/\sigma^2)$ ενώ η ενδιαφέρουσα παράμετρος είναι το $\theta/\delta = m$.

2.4.2 Εκτιμητές με συρρικνωτή

Οι θεωρούμενοι εδώ εκτιμητές είναι, όπως στο Κεφ.1, με συρρικνωτή "*matriciel*", δηλαδή δεν εφαρμόζουν τον ίδιο συντελεστή συρρίκνωσης στις διάφορες συνιστώσες του εκτιμητή ελαχίστων τετραγώνων ϕ_0.

Έστω b μια θετική, συμμετρική, διγραμμική μορφή επί του Θ, τάξεως μεγαλύτερης ή ίσης του 3 και c ένας ενδομορφισμός του Θ. Έστω h μια συνάρτηση από τον $(\mathbb{R}_+)^2$ στον $\mathbb{R}_+$, διαφορίσιμη. Επιπλέον, υποθέτουμε οτι υπάρχει ένα ανοικτό, μη-κενό υποσύνολο του $(\mathbb{R}_+)^2$, στο οποίο η μερική παράγωγος της h αναφορικά με την πρώτη μεταβλητή, την οποία σημειώνουμε

h_1', είναι συνεχής και όχι ταυτοτικά μηδέν. Ο συνήθης αμερόληπτος εκτιμητής του σ^2 είναι ο $s^2(y) = \|y - \phi_0(y)\|^2/(n-k)$. Ορίζουμε τον *εκτιμητή με συρρικνωτή* που συνδέεται με τις b, c, h από τη σχέση:

$$\phi(y) = \Big\{ id_\Theta - h\Big(\bar{b}(\phi_0(y)), s^2(y)\Big)c\Big\}\phi_0(y) \tag{2.4}$$

όπου id_Θ είναι ο ταυτοτικός ενδομορφισμός του Θ.

2.4.3 Αναγκαία συνθήκη για την παραδοχή

Εξετάζοντας τα αποτελέσματα της Πρότασης 2.3.1 για τους εκτιμητές της μορφής (2.4) προκύπτει:

Πρόταση 2.4.2
Αν ένας εκτιμητής της μορφής (2.4) είναι παραδεκτός, θεωρώντας οτι το σ^2 παίρνει τιμές σε ένα συμπαγές διάστημα του $\mathbb{R}_+^*$, τότε υπάρχει ένα $\varrho \in \mathbb{R}^*$ τέτοιο ώστε $c = \varrho v_\Theta b$, όπου v_Θ^{-1} ο περιορισμός του v^{-1} στο Θ.

Απόδειξη
α) Έστω ϕ ένας παραδεκτός εκτιμητής και έστω f μια C^∞-συνάρτηση από τον $\Theta \times \mathbb{R}^*$ στον $\mathbb{R}$, τέτοια ώστε:

$$\phi(y) = \frac{\frac{\partial f}{\partial u}(\phi_0(y), t(y))}{\frac{\partial f}{\partial z}(\phi_0(y), t(y))} \qquad (\nu\text{- } \sigma.\pi.)$$

Έστω $\mathcal{E} = (e_1, \ \dots \ , e_n)$ μια βάση v^{-1}-ορθοκανονική του E, τέτοια ώστε οι k πρώτοι όροι, $(e_1, \ \dots \ , e_k)$, να αποτελούν μια βάση του Θ, b-ορθογώνια. Αν $y = \sum_{i=1}^n y_i e_i$ τότε:

$$\phi_0(y) = \sum_{i=1}^k y_i e_i \quad , \quad s^2(y) = \frac{1}{n-k}\sum_{i=k+1}^n y_i^2 \quad \text{και} \quad \bar{b}\Big(\phi_0(y)\Big) = \sum_{i=1}^k b_i y_i^2$$

όπου, λόγω της υπόθεσης $rg(b) \geq 3$, τουλάχιστον τρείς από τους συντελεστές b_i είναι μη μηδενικοί.

Έστω (c_{ij}) ο πίνακας του ενδομορφισμού c σε αυτή τη βάση. Η Πρόταση θα αποδειχθεί, δείχνοντας οτι, για κάθε i διάφορο του j, $c_{ij} = 0$ και οτι υπάρχει ένα ϱ τέτοιο ώστε, για κάθε i, $c_{ii} = \varrho b_i$(σημειώνουμε οτι το ϱ δεν μπορεί να είναι μηδέν, γιατί σε αυτή την περίπτωση ο ϕ θα ταυτιζόταν με τον ϕ_0, ο οποίος όμως δεν είναι παραδεκτός σύμφωνα με τον $Stein$[1956]).

β) Θέτοντας,

$$g(y_1, \ \dots \ , y_k, s^2(y)) = f\Big(\sum_{i=1}^k y_i e_i, -\frac{1}{2}(\sum_{i=1}^k y_i^2 + (n-k)s^2(y))\Big) ,$$

για $1 \leq i \leq k$, έχουμε:

$$\begin{aligned} \frac{\partial g}{\partial y_i}(y_1, \ \dots \ , y_k, s^2(y)) &= \frac{\partial f}{\partial u}\Big(\sum_{i=1}^k y_i e_i, -\frac{1}{2}(\sum_{i=1}^k y_i^2 + (n-k)s^2(y))\Big) . e_i \\ &- \frac{\partial f}{\partial z}\Big(\sum_{i=1}^k y_i e_i, -\frac{1}{2}(\sum_{i=1}^k y_i^2 + (n-k)s^2(y))\Big) y_i \end{aligned}$$

Άρα,

$$\sum_{i=1}^{k} \frac{\partial g}{\partial y_i}(y_1, \ \dots \ , y_k, s^2(y))e_i = \frac{\partial f}{\partial u}\Big(\sum_{i=1}^{k} y_i e_i, t(y)\Big) - \frac{\partial f}{\partial z}\Big(\sum_{i=1}^{k} y_i e_i, t(y)\Big) \sum_{i=1}^{k} y_i e_i$$

και

$$\frac{\partial g}{\partial s^2}(y_1, \ \dots \ , y_k, s^2(y)) = -\frac{n-k}{2} \ \frac{\partial f}{\partial z}\Big(\sum_{i=1}^{k} y_i e_i, t(y)\Big)$$

απ΄ όπου προκύπτει,

$$\begin{aligned}
\phi(y) &= \frac{\frac{\partial f}{\partial u}(\phi_0(y), t(y))}{\frac{\partial f}{\partial z}(\phi_0(y), t(y))} \\
&= \frac{\sum_{i=1}^{k} \frac{\partial g}{\partial y_i}(y_1, \ \dots \ , y_k, s^2(y))e_i + \sum_{i=1}^{k} y_i e_i \frac{\partial f}{\partial z}\Big(\sum_{i=1}^{k} y_i e_i, t(y)\Big)}{-\frac{2}{n-k} \ \frac{\partial g}{\partial s^2}(y_1, \ \dots \ , y_k, s^2(y))} \\
&= -\frac{n-k}{2} \ \frac{\sum_{i=1}^{k} \frac{\partial g}{\partial y_i}(y_1, \ \dots \ , y_k, s^2(y))e_i}{\frac{\partial g}{\partial s^2}(y_1, \ \dots \ , y_k, s^2(y))} + \frac{\sum_{i=1}^{k} y_i e_i \frac{\partial f}{\partial z}\Big(\sum_{i=1}^{k} y_i e_i, t(y)\Big)}{-\frac{2}{n-k} \ \frac{\partial g}{\partial s^2}(y_1, \ \dots \ , y_k, s^2(y))} \\
&= \sum_{i=1}^{k} y_i e_i - \frac{n-k}{2} \sum_{i=1}^{k} \frac{\frac{\partial g}{\partial y_i}(y_1, \ \dots \ , y_k, s^2(y))}{\frac{\partial g}{\partial s^2}(y_1, \ \dots \ , y_k, s^2(y))} e_i \qquad (\nu - \sigma.\pi.)
\end{aligned}$$

Για να είναι, λοιπόν, ένας εκτιμητής της μορφής (2.4) παραδεκτός είναι αναγκαίο να ισχύει οτι, για κάθε i με $1 \leq i \leq k$,

$$h\Big(\sum_{j=1}^{k} b_j y_j^2, s^2(y)\Big) \sum_{j=1}^{k} c_{ij} y_j \overset{\sigma.\pi.}{=} \frac{n-k}{2} \ \frac{\partial g}{\partial y_i}(y_1, \ \dots \ , y_k, s^2(y))\Big(\frac{\partial g}{\partial s^2}(y_1, \ \dots \ , y_k, s^2(y))\Big)^{-1} \quad (2.5)$$

αφού, $\quad c\Big(\phi_0(y)\Big) = \sum_{i=1}^{k} y_i c(e_i) = \sum_{i=1}^{k} y_i \Big(\sum_{j=1}^{k} c_{ji} e_j\Big) = \sum_{j=1}^{k} \Big(\sum_{i=1}^{k} c_{ji} y_i\Big) e_j = \sum_{i=1}^{k} \Big(\sum_{j=1}^{k} c_{ij} y_j\Big) e_i$

Μια άμεση συνέπεια της (2.5) είναι οτι, για σχεδόν κάθε y,

$$\frac{\partial g}{\partial y_i}(y_1, \ \dots \ , y_k, s^2(y))\Big(\sum_{j=1}^{k} c_{ij} y_j\Big)^{-1}$$

είναι ανεξάρτητη του i. Προκύπτει, λοιπόν, οτι: $\forall (i,p) \in \{1, \ \dots \ , k\}^2$,

$$\frac{\partial^2 g}{\partial s^2 \partial y_i}(y_1, \ \dots \ , y_k, s^2(y)) \sum_{j=1}^{k} c_{pj} y_j \overset{\sigma.\pi.}{=} \frac{\partial^2 g}{\partial s^2 \partial y_p}(y_1, \ \dots \ , y_k, s^2(y)) \sum_{j=1}^{k} c_{ij} y_j \qquad (2.6)$$

Επιπλέον, παρατηρούμε οτι η (2.5) μας δίνει, για κάθε i, μία έκφραση της $\partial g/\partial y_i$, συγκεκριμένα

$$\frac{n-k}{2} \ \frac{\partial g}{\partial y_i}(y_1, \ \dots \ , y_k, s^2(y)) \overset{\sigma.\pi.}{=} \frac{\partial g}{\partial s^2}(y_1, \ \dots \ , y_k, s^2(y)) h\Big(\sum_{j=1}^{k} b_j y_j^2, s^2(y)\Big) \sum_{j=1}^{k} c_{ij} y_j$$

Παραγωγίζοντας τη σχέση αυτή, λαμβάνουμε:

$$
\begin{aligned}
&\frac{n-k}{2}\ \frac{\partial^2 g}{\partial y_p \partial y_i}(y_1,\ \dots\ ,y_k,s^2(y))\\
&\overset{\sigma.\pi.}{=} h_1'\Big(\sum_{j=1}^{k} b_j y_j^2, s^2(y)\Big)2b_p y_p\Big(\sum_{j=1}^{k} c_{ij}y_j\Big)\frac{\partial g}{\partial s^2}(y_1,\ \dots\ ,y_k,s^2(y))\\
&\quad +h\Big(\sum_{j=1}^{k} b_j y_j^2, s^2(y)\Big)c_{ip}\frac{\partial g}{\partial s^2}(y_1,\ \dots\ ,y_k,s^2(y))\\
&\quad +h\Big(\sum_{j=1}^{k} b_j y_j^2, s^2(y)\Big)\Big(\sum_{j=1}^{k} c_{ij}y_j\Big)\frac{\partial^2 g}{\partial y_p \partial s^2}(y_1,\ \dots\ ,y_k,s^2(y))\\
&= \Big[2h_1'\Big(\bar{b}(\phi_0(y)),s^2(y)\Big)b_p y_p \sum_{j=1}^{k} c_{ij}y_j + h\Big(\bar{b}(\phi_0(y)),s^2(y)\Big)c_{ip}\Big]\frac{\partial g}{\partial s^2}(y_1,\ \dots\ ,y_k,s^2(y))\\
&\quad +\frac{\partial^2 g}{\partial s^2 \partial y_p}(y_1,\ \dots\ ,y_k,s^2(y))h\Big(\bar{b}(\phi_0(y)),s^2(y)\Big)\sum_{j=1}^{k} c_{ij}y_j
\end{aligned}
$$

Ακριβώς με τον ίδιο τρόπο, λαμβάνουμε,

$$
\begin{aligned}
&\frac{n-k}{2}\ \frac{\partial^2 g}{\partial y_i \partial y_p}(y_1,\ \dots\ ,y_k,s^2(y))\\
&= \Big[2h_1'\Big(\bar{b}(\phi_0(y)),s^2(y)\Big)b_i y_i \sum_{j=1}^{k} c_{pj}y_j + h\Big(\bar{b}(\phi_0(y)),s^2(y)\Big)c_{pi}\Big]\frac{\partial g}{\partial s^2}(y_1,\ \dots\ ,y_k,s^2(y))\\
&\quad +\frac{\partial^2 g}{\partial s^2 \partial y_i}(y_1,\ \dots\ ,y_k,s^2(y))h\Big(\bar{b}(\phi_0(y)),s^2(y)\Big)\sum_{j=1}^{k} c_{pj}y_j
\end{aligned}
$$

Χρησιμοποιώντας την (2.6) και εξισώνοντας τις δύο παραπάνω εκφράσεις παίρνουμε:

$$
\begin{aligned}
&2h_1'\Big(\bar{b}(\phi_0(y)),s^2(y)\Big)b_p y_p \sum_{j=1}^{k} c_{ij}y_j + h\Big(\bar{b}(\phi_0(y)),s^2(y)\Big)c_{ip}\\
&= 2h_1'\Big(\bar{b}(\phi_0(y)),s^2(y)\Big)b_i y_i \sum_{j=1}^{k} c_{pj}y_j + h\Big(\bar{b}(\phi_0(y)),s^2(y)\Big)c_{pi}
\end{aligned}
$$

απ᾽ όπου, θέτοντας

$$P_{ip}(y_1,\ \dots\ ,y_k) = b_p y_p \sum_{j=1}^{k} c_{ij}y_j$$

προκύπτει οτι, για κάθε $(i,p) \in \{1,\ \dots\ ,k\}^2$ και για σχεδόν κάθε y,

$$2h_1'\Big(\bar{b}(\phi_0(y)),s^2(y)\Big)\Big[P_{ip}(y_1,\ \dots\ ,y_k) - P_{pi}(y_1,\ \dots\ ,y_k)\Big] = h\Big(\bar{b}(\phi_0(y)),s^2(y)\Big)(c_{pi} - c_{ip}) \quad (2.7)$$

Από την υπόθεση(βλέπε παράγραφο 2.4.2), υπάρχει ένα ανοικτό, μη κενό σύνολο $A \subset (\mathbb{R}_+)^2$ στο οποίο η h_1' είναι συνεχής και όχι ταυτοτικά μηδέν. Αν θεωρήσουμε τη διανυσματική συνάρτηση $F : E \to (\mathbb{R}_+)^2$ με $F(y) = \Big(\bar{b}(\phi_0(y)), s^2(y)\Big)$, αυτή είναι συνεχής άρα η αντίστροφη εικόνα του A, $F^{-1}(A)$, είναι ανοικτό υποσύνολο του E. Δηλαδή, στο ανοικτό $F^{-1}(A)$ $(\subset E)$ η $h_1' \circ F$ είναι συνεχής και όχι ταυτοτικά μηδέν, άρα υπάρχει ένα μη-κενό, ανοικτό $G \subset F^{-1}(A)$ ώστε $\forall y \in G$, $h_1' \circ F(y) \neq 0$ και να ικανοποιούνται οι ισότητες της μορφής (2.7).

Προκύπτει, λοιπόν, οτι για κάθε τριάδα (i, p, q) έχουμε την ισότητα των πολυωνύμων:

$$(c_{qi} - c_{iq})(P_{ip} - P_{pi}) = (c_{pi} - c_{ip})(P_{iq} - P_{qi}) \tag{2.8}$$

Πράγματι, από την (2.7) για τα ζεύγη (i, p) και (i, q) ξεχωριστά, παίρνουμε:

$$2h_1'\Big(\bar{b}(\phi_0(y)), s^2(y)\Big)\Big[P_{ip}(y_1, \ \ldots \ , y_k) - P_{pi}(y_1, \ \ldots \ , y_k)\Big] = h\Big(\bar{b}(\phi_0(y)), s^2(y)\Big)(c_{pi} - c_{ip})$$

και

$$2h_1'\Big(\bar{b}(\phi_0(y)), s^2(y)\Big)\Big[P_{iq}(y_1, \ \ldots \ , y_k) - P_{qi}(y_1, \ \ldots \ , y_k)\Big] = h\Big(\bar{b}(\phi_0(y)), s^2(y)\Big)(c_{qi} - c_{iq})$$

απ΄ όπου διαιρώντας κατά μέλη λαμβάνουμε την (2.8).

Ταυτοποιώντας, από τα πολυώνυμα της (2.8), συμπεραίνουμε οτι, για κάθε τριάδα (i, p, q) διαφορετικών στοιχείων έχουμε:

$$(c_{pi} - c_{ip})b_q c_{iq} = 0 \quad \text{και} \quad (c_{qi} - c_{iq})b_p c_{iq} = (c_{pi} - c_{ip})b_q c_{ip} \tag{2.9}$$

(οι σχέσεις αυτές προέκυψαν εξισώνοντας τους συντελεστές του y_q^2 και του $y_q y_p$ αντιστοίχως στην (2.8))

Επίσης, για κάθε ζεύγος (i, p), υπάρχει ένα q, διάφορο των i και p, τέτοιο ώστε $b_q \neq 0$ (γιατί έχουμε υποθέσει οτι $rg(b) \geq 3$). Προκύπτει, λοιπόν, από την (2.9) και υποθέτοντας οτι $c_{ip} \neq c_{pi}$, οτι $c_{iq} = 0$ και συνεπώς $c_{ip} = 0$. Όμως, επειδή τα i και p έχουν συμμετρικούς ρόλους (δηλ. η (2.9) ισχύει και αν εναλλάξουμε τα i και p), η υπόθεση αυτή μας δίνει ακριβώς όμοια οτι και $c_{pi} = 0$. Έχουμε λοιπόν αναγκαστικά οτι $c_{ip} = c_{pi}$.

Επειδή οι ισότητες (2.7) ικανοποιούνται σε όλο το G, προκύπτει λοιπόν, από τη μη-μηδενικότητα της $h_1'\Big(\bar{b}(\phi_0(y)), s^2(y)\Big)$ στο G, η ισότητα των πολυωνύμων P_{ip} και P_{pi}, για κάθε ζεύγος (i, p). Δηλαδή,

$$b_p y_p \sum_{j=1}^{k} c_{ij} y_j = b_i y_i \sum_{j=1}^{k} c_{pj} y_j$$

Ταυτοποιώντας τους όρους του πολυωνύμου, συμπεραίνουμε οτι για κάθε τριάδα (i, j, p) με $i \neq j \neq p$ έχουμε $b_p c_{ij} = 0$ (εξισώνοντας τους συντελεστές του $y_p y_j$) και για κάθε ζεύγος (i, p), $b_p c_{ii} = b_i c_{pp}$ (συντελεστές του $y_i y_p$). Χρησιμοποιώντας, εκ νέου, την υπόθεση οτι τουλάχιστον τρείς από τους συντελεστές b_p είναι μη-μηδενικοί, λαμβάνουμε οτι:

$$\forall i \neq j \ \ , \quad c_{ij} = 0 \quad (\text{αφού} \ \ b_p c_{ij} = 0 \quad \text{και} \quad b_p \neq 0)$$

και οτι, υπάρχει $\varrho \in \mathbb{R}^*$, τέτοιο ώστε, για κάθε i, $c_{ii} = \varrho b_i$.

Πράγματι, επειδή για κάθε i, $c_{ii} = b_i c_{pp}/b_p$ έχουμε:

$$\sum_{i=1}^{k} c_{ii} = \frac{c_{pp}}{b_p} \sum_{i=1}^{k} b_i \quad \text{δηλ.} \quad \frac{c_{pp}}{b_p} = \frac{tr\, c}{\sum_{i=1}^{k} b_i}$$

Λαμβάνοντας λοιπόν $\varrho = \dfrac{tr\, c}{\sum_{i=1}^{k} b_i}$ έχουμε το ζητούμενο.

□

2.4.4 Πλήρης Υποκλάση

Θεωρούμε τους εκτιμητές της μορφής (2.4) με $c = v_\Theta b$, διότι μόνο αυτοί μπορούν να είναι παραδεκτοί (μπορούμε πάντοτε να θεωρούμε την περίπτωση όπου $\varrho = 1$, αντικαθιστώντας την h με ϱh). Μεταξύ των εκτιμητών της μορφής (2.4) αυτοί για τους οποίους $c = v_\Theta b$ σχηματίζουν μια πλήρη κλάση.

Πρόταση 2.4.3

Μια ικανή συνθήκη ώστε ο εκτιμητής

$$\phi(y) = \Big\{id_\Theta - h\Big(\bar{b}(\phi_0(y)), s^2(y)\Big)v_\Theta b\Big\}\phi_0(y)$$

όπου $rg(b) \geq 3$, να κυριαρχεί τον εκτιμητή ελαχίστων τετραγώνων, ϕ_0, είναι να υπάρχουν δύο πραγματικοί αριθμοί l_1 και l_2 τέτοιοι ώστε:

$i)\forall u > 0,\ t \mapsto t^{l_1}h(t,u)$ να είναι αύξουσα συνάρτηση

$ii)\forall t > 0,\ u \mapsto u^{l_2}h(t,u)$ να είναι φθίνουσα συνάρτηση και

$iii)$η $(t/u)h(t,u)$ να είναι ομοιόμορφα φραγμένη από το

$$2\left(\frac{tr(v_\Theta q v_\Theta b)}{pgvp(v_\Theta q v_\Theta b)} - 2l_1\right)\frac{n-k}{n-k-2-4l_2}$$

όπου $pgvp(v_\Theta q v_\Theta b)$ σημειώνει τη μεγαλύτερη ιδιοτιμή του ενδομορφισμού $v_\Theta q v_\Theta b$.

Παρατήρηση 2.4.1

Οι υποθέσεις της Προτάσεως 2.4.3 είναι συμβατές αν $l_2 < -1$ διότι, διαφορετικά θα είχαμε:

$$\forall t > 0,\ \lim_{u\to 0} u^{-1}h(t,u) = \lim_{u\to 0}\Big(u^{-1-l_2}u^{l_2}h(t,u)\Big) = +\infty$$

το οποίο αντιφάσκει στον ισχυρισμό οτι η $tu^{-1}h(t,u)$ είναι φραγμένη.

Απόδειξη

Η Πρόταση αυτή αποτελεί μια ειδική περίπτωση της Πρότασης 1.5.6, όπου εδώ $c = v_\Theta b$. Συμπεραίνεται λοιπόν από αυτή επαληθεύοντας τις συνθήκες της. Αν $\widetilde{h}(t,u)$ είναι η συνάρτηση συρρίκνωσης εκείνης της Πρότασης τότε $\widetilde{h}(t,u) = h(t, 1/u)$. Οπότε,
$\forall u > 0,\ t \mapsto t^{l_1}\widetilde{h}(t,u) = t^{l_1}h(t,1/u)$ είναι αύξουσα συνάρτηση (σύμφωνα με το $i)$)
$\forall t > 0,\ u \mapsto u^{-l_2}\widetilde{h}(t,u) = u^{-l_2}h(t,1/u)$ είναι αύξουσα συνάρτηση (σύμφωνα με το $ii)$) και
η $tu\widetilde{h}(t,u)$= $tuh(t,1/u)$ είναι ομοιόμορφα φραγμένη από το

$$2\left(\frac{tr(v_\Theta q v_\Theta b)}{pgvp(v_\Theta q v_\Theta b)} - 2l_1\right)\frac{n-k}{n-k-2-4l_2}$$

(σύμφωνα με το $iii)$). Όμως αυτή η σταθερά είναι ίση με την R_2 της Πρότασης 1.5.6, αφού τα φάσματα των ενδομορφισμών ${}^tcqcb^-$ και $v_\Theta q v_\Theta b$ συμπίπτουν και επομένως

$$\lambda_M = \sigma_M = pgvp(v_\Theta q v_\Theta b)$$

απ΄ όπου εξασφαλίζεται άμεσα και η υπόθεση $\lambda_m \geq 0$, αφού $\lambda_m = \sigma_m \geq 0$ (βλέπε Κεφάλαιο 1, παράγραφος 1.3.1).

□

Παρατηρούμε οτι η επιλογή του b που μεγιστοποιεί το παραπάνω φράγμα είναι $b = (v_\Theta q v_\Theta)^{-1}$, αφού

$$\begin{aligned} 2\left(\frac{tr(v_\Theta q v_\Theta b)}{pgvp(v_\Theta q v_\Theta b)} - 2l_1\right)\frac{n-k}{n-k-2-4l_2} &\leq 2\left(\frac{pgvp(v_\Theta q v_\Theta b)k}{pgvp(v_\Theta q v_\Theta b)} - 2l_1\right)\frac{n-k}{n-k-2-4l_2} \\ &= \frac{2(k-2l_1)(n-k)}{n-k-2-4l_2} \end{aligned}$$

φράγμα το οποίο λαμβάνεται όταν το $b = (v_\Theta q v_\Theta)^{-1}$ (αυτό δεν αντιστοιχεί αναγκαίως στο μέγιστο κέρδος, αναφορικά με τον εκτιμητή ελαχίστων τετραγώνων).

Προκύπτει, λοιπόν, οτι με κάθε συνάρτηση $h(t,u)$ που ικανοποιεί τις συνθήκες $i)$ και $ii)$ και επίσης

$$\sup_{t,u} \frac{t}{u} h(t,u) \leq \frac{2(k-2l_1)(n-k)}{n-k-2-4l_2}$$

μπορούμε να συνδέσουμε έναν εκτιμητή με συρρικνωτή κυριαρχώντας ομοιομόρφως τον ϕ_0, αναφορικά με κάθε τετραγωνική συνάρτηση απωλείας. Ένας τέτοιος εκτιμητής θα είχε τότε την μορφή:

$$\phi(y) = \left\{ id_\Theta - h\Big(\bar{b}'(\phi_0(y)), s^2(y)\Big) v_\Theta b' \right\} \phi_0(y)$$

με $b' = (v_\Theta q v_\Theta)^{-1}$ και $v_\Theta b' = (v_\Theta q)^{-1}$.

Κεφάλαιο 3

ΙΚΑΝΗ ΚΑΙ ΑΝΑΓΚΑΙΑ ΣΥΝΘΗΚΗ ΓΙΑ ΤΗΝ ΚΥΡΙΑΡΧΙΑ ΤΟΥ ΣΥΝΗΘΟΥΣ ΕΚΤΙΜΗΤΗ ΕΛΑΧΙΣΤΩΝ ΤΕΤΡΑΓΩΝΩΝ

3.1

Καθώς οι *James* και *Stein*[1961] θέτουν τις βάσεις της θεωρίας των εκτιμητών με συρρικνωτή, πληθώρα συγγραφέων μελετά τις πιθανές "βελτιώσεις" του εκτιμητή ελαχίστων τετραγώνων, υπό διάφορες οπτικές (βλ. π.χ.*Judge* και *Bock*[1978] ή *Berger*[1982]).

Η πλειοψηφία των συγγραφέων θεωρεί τετραγωνικές συναρτήσεις απώλειας, οι *Brardwein* και *Strawderman*[1980] μελετούν κυρτές συναρτήσεις απώλειας, για το σύνολο των νόμων με σφαιρική συμμετρία. Από την άλλη, κάποιοι συγγραφείς συγκρίνουν τους εκτιμητές τους με κριτήριο την τάξη των διγραμμικών μορφών, "δεύτερη ροπή της διανυσματικής τυχαίας μεταβλητής:«εκτίμηση μείον την αληθινή τιμή»" (αυτό αποτελεί το κριτήριο του μέσου τετραγωνικού σφάλματος(MSE)) ή ισοδύναμα αναφορικά με το σύνολο των τετραγωνικών κινδύνων. Αυτή η ιδιότητα είναι ενδιαφέρουσα επειδή δείχνει τη "ρωμαλεότητα" του εκτιμητή, κυριαρχώντας με την έννοια του MSE.

Ο *Brown*[1975], για την κλάση των εκτιμητών της μορφής:

$$\phi(y) = \phi_0(y) - h\Big(\|\phi_0(y)\|^2\Big)\phi_0(y)$$

αποδεικνύει ότι δεν υπάρχει εκτιμητής που να βελτιώνει τον ε.ε.τ. ϕ_0 για το σύνολο των τετραγωνικών απωλειών. Το αποτέλεσμα αυτό αποτελεί πόρισμα μιας ικανής και αναγκαίας συνθήκης κυριαρχίας επί του ε.ε.τ. για ένα σύνολο τετραγωνικών συναρτήσεων απώλειας. Εδώ θα ασχοληθούμε, σύμφωνα με τον *C.Robert*[1987] ([41]), με τη γενίκευση εκείνης της συνθήκης για την κλάση των εκτιμητών με συρρικνωτή "*matriciel*" του Κεφ.2, λαμβάνοντας αρχικά $b = v_\Theta^{-1}$. Στη συνέχεια γενικεύουμε το αποτέλεσμα για οποιαδήποτε θετική, συμμετρική, διγραμμική μορφή b.

3.2 ΤΟ ΜΟΝΤΕΛΟ

Έστω E διανυσματικός χώρος, διάστασης n, επί του $\mathbb{R}$, Θ ένας διανυσματικός υπόχωρος του E, διάστασης k ($3 \leq k < n-1$). Παρατηρούμε $y \in E$, y είναι η πραγματοποίηση μίας κανονικής τ.μ μέσου $\theta \in \Theta$ και διασποράς $\sigma^2 v$, όπου v είναι μια γνωστή συμμετρική διγραμμική μορφή, θετικά ορισμένη επί του E^*(δυϊκός του E). Δεχόμαστε και εδώ ότι η διασπορά είναι γνωστή εκτός από τον πολλαπλασιαστικό παράγοντα σ^2.

Οι *εκτιμητές με συρρικνωτή* του θ, τους οποίους θεωρούμε, είναι της μορφής:

$$\phi(y) = \Big(id_\Theta - h\Big(\overline{v_\Theta^{-1}}(\phi_0(y)), s^2(y)\Big)c\Big)\phi_0(y) \tag{3.1}$$

όπου:

- id_Θ είναι ο ταυτοτικός ενδομορφισμός του Θ,

- c είναι ένας ενδομορφισμός του Θ,

- $\phi_0(y)$ είναι η v^{-1}-ορθογώνια προβολή επί του Θ, δηλ. ο ε.ε.τ του θ και $s^2(y) = \frac{1}{n-k}\overline{v^{-1}}\Big(y - \phi_0(y)\Big)$ ο συνήθης αμερόληπτος εκτιμητής του σ^2,

- h είναι μία μετρήσιμη απεικόνιση από τον $\mathbb{R}_+ \times \mathbb{R}_+$ στον $\mathbb{R}$.

Παρατηρούμε ότι, αναφορικά με την κλάση των εκτιμητών (2.4) του Κεφ.2, εδώ θεωρούμε την ειδική περίπτωση όπου η τετραγωνική μορφή $\overline{b}$, που υπεισέρχεται στον ορισμό της h, είναι η $\overline{v_\Theta^{-1}}$.

Δεδομένης μιας θετικής, συμμετρικής, διγραμμικής μορφής, q, επί του Θ, η συνάρτηση απώλειας που συνδέεται με αυτήν ορίζεται να είναι η:

$$\frac{1}{\sigma^2}\,\overline{q}(\widehat{\theta} - \theta)$$

αν $\widehat{\theta}$ είναι ο θεωρούμενος εκτιμητής του θ. Εδώ θα θεωρήσουμε μια κλάση συμμετρικών, διγραμμικών μορφών Q. Θα λέμε ότι ένας εκτιμητής ϕ *κυριαρχεί τον* ϕ_0 *ομοιομόρφως στο* (θ, σ^2) *επί του* Q αν, για κάθε $(\theta, \sigma^2) \in \Theta \times \mathbb{R}_+^*$ και για κάθε $q \in Q$, έχουμε:

$$\frac{1}{\sigma^2}\,\mathbb{E}_{\theta,\sigma}\Big[\overline{q}(\phi(y) - \theta)\Big] \leq \frac{1}{\sigma^2}\,\mathbb{E}_{\theta,\sigma}\Big[\overline{q}(\phi_0(y) - \theta)\Big] \ ,$$

όπου για κάθε g μετρήσιμη απεικόνιση, σημειώνουμε

$$\mathbb{E}_{\theta,\sigma}\Big[g(y)\Big] = \int_E g(y)P_{\theta,\sigma}(dy) \ \ , \quad P_{\theta,\sigma} = N(\theta, \sigma^2 v)$$

Θα αποδείξουμε παρακάτω μια ικανή και αναγκαία συνθήκη για την ύπαρξη ενός εκτιμητή ϕ κυριαρχώντας τον ϕ_0 ομοιομόρφως στο (θ, σ^2) επί του Q. Το αποτέλεσμα του $Brown$[1975], που αφορά την ειδική παρίπτωση όπου το σ^2 είναι γνωστό και ο $c = id_\Theta$ φαίνεται τότε σαν πόρισμα του θεωρήματος αυτού.

3.3 ΙΚΑΝΗ ΚΑΙ ΑΝΑΓΚΑΙΑ ΣΥΝΘΗΚΗ

Θεωρώντας μια οικογένεια συναρτήσεων απωλείας Q, αν για κάθε $q \in Q$, $tr(v_\Theta qc) > 0$, έχουμε:

Θεώρημα 3.3.1

Υπάρχει ένας εκτιμητής ϕ *της μορφής (3.1) που κυριαρχεί τον* ϕ_0 *ομοιομόρφως στο* (θ, σ^2) *επί του* Q *αν και μόνον αν*

$$\inf_{q\in Q}\Big\{tr(v_\Theta qc) - 2pgvp(v_\Theta qc)\Big\} > 0 \tag{3.2}$$

Πρέπει να σημειώσουμε ότι η συνθήκη (3.2) είναι λιγότερο περιοριστική από την αντίστοιχη του $Brown$:

$$\inf_{q\in Q}\Big\{tr(v_\Theta q)-2pgvp(v_\Theta q)\Big\}>0\ \ .$$

Η εισαγωγή, λοιπόν, του ενδομορφισμού συρρίκνωσης, c, μας επιτρέπει, για κάποιες κλάσεις Q, να έχουμε ομοιόμορφη κυριαρχία η οποία δεν είναι εφικτή για τους εκτιμητές με συρρικνωτή κλίμακας (ιδιαιτέρως στις περιπτώσεις της "*multicolinéarité*").

Η έκφραση $\{tr(v_\Theta qc)-2pgvp(v_\Theta qc)\}$, εμφανίζεται συχνά στις συνθήκες άνω φράγματος της συναρτήσεως συρρίκνωσης (όπως στους $Judge$ και $Bock$[1978], $Cellier$ και $Fourdrinier$[1985] και στο Κεφ.1). Όταν αυτό το φράγμα ήταν αρνητικό υπήρχε αδυναμία να προταθούν εκτιμητές καλύτεροι από τον ϕ_0. Το παραπάνω όμως Θεώρημα αποδεικνύει ότι κάτι τέτοιο δεν μπορεί να γίνει. Σημειώνουμε ακόμη ότι η συνθήκη $tr(v_\Theta qc)-2pgvp(v_\Theta qc)>0$ εμφανίζεται ήδη σαν αναγκαία στις περιπτώσεις που η συνάρτηση συρρίκνωσης h είναι της μορφής $h(t,u)=\alpha t^{\lambda_1}u^{\lambda_2}$ (βλέπε Κεφ.1).

Να σημειώσουμε ότι αυτό το κεφάλαιο δεν περιλαμβάνει τη μελέτη εκτιμητών που εξαρτώνται από την επιλογή του κινδύνου, όπως, για παράδειγμα θα ήταν η περίπτωση που $c_q=(v_\Theta q)^{-1}$, ο οποίος μεγιστοποιεί την ποσότητα $\{tr(v_\Theta qc)-2pgvp(v_\Theta qc)\}$ (σε έναν βαθμωτό παράγοντα).

Απόδειξη

Το ότι η συνθήκη (3.2) είναι ικανή, για την ύπαρξη ενός εκτιμητή ϕ που κυριαρχεί τον ϕ_0 ομοιομόρφως στο (θ,σ^2) επί του Q, εξασφαλίζεται από την Πρόταση 1.5.6(Κεφ.1), παίρνοντας $\mu_1=1$, οπότε $R_2>0$.

Για να αποδείξουμε ότι η συνθήκη αυτή είναι και αναγκαία θα χρησιμοποιήσουμε την τεχνική του $Brown$[1975] που στηρίζεται, όπως εκείνη του $Stein$[1956], στην ανισότητα $Cramer-Rao$.

Για να κυριαρχεί ένας εκτιμητής ϕ τον ϕ_0, ομοιομόρφως στο (θ,σ^2) επί του Q, θα πρέπει, για κάθε $q\in Q$ και για κάθε $(\theta,\sigma^2)\in\Theta\times\mathbb{R}_+^*$, να ισχύει:

$$\mathbb{E}_{\theta,\sigma}\Big[\overline{q}(\phi(y)-\theta)\Big]\leq\mathbb{E}_{\theta,\sigma}\Big[\overline{q}(\phi_0(y)-\theta)\Big]=\sigma^2tr(v_\Theta q)$$

Έστω $(e_1,\ \dots\ ,e_k)$ μια βάση v_Θ^{-1}-ορθοκανονική του Θ και q-ορθογώνια. Λοιπόν, σε αυτή τη βάση, έχουμε:

$$\theta=\sum_{i=1}^k\theta_ie_i\quad\text{ και }\quad\mathbb{E}_{\theta,\sigma}\Big[(\phi(y)-\theta)\Big]=b_\sigma(\theta)=\sum_{i=1}^kb_{\sigma_i}(\theta)e_i\ \text{ το διάνυσμα της μεροληψίας.}$$

Για κάθε j, σημειώνουμε $(\partial_jb_{\sigma_i})(\theta)$ τη μερική παράγωγο του b_{σ_i} αναφορικά με την j-στή συντεταγμένη του θ. Για κάθε i, η ανισότητα των $Cramer-Rao$:

$$\sum_{i=1}^k\mathbb{E}_{\theta,\sigma}\Big(\phi_i(y)-\theta_i\Big)^2\geq\sum_{i=1}^k\Big(b_{\sigma_i}(\theta)\Big)^2+k\sigma^2+2\sigma^2\sum_{i=1}^k\partial_ib_{\sigma_i}(\theta)+\sigma^2\sum_{i=1}^k\sum_{j=1}^k\Big(\partial_jb_{\sigma_i}(\theta)\Big)^2$$

εφαρμοζόμενη στο μοντέλο όπου $\theta_j(j\neq i)$ και σ^2 θεωρούνται γνωστά, και όπου μόνο το θ_i είναι άγνωστο, παίρνει τη μορφή:

$$\mathbb{E}_{\theta,\sigma}\Big[(\phi_i(y)-\theta_i)^2\Big]\geq\Big(b_{\sigma_i}(\theta)\Big)^2+\Big(1+\partial_ib_{\sigma_i}(\theta)\Big)^2\sigma^2$$

Σημεώνουμε $(q_i)_{1\leq i\leq k}$, τις ιδιοτιμές του $v_\Theta q$. Για να κυριαρχεί ο ϕ τον ϕ_0, θα πρέπει λοιπόν

$$\sum_{i=1}^{k} q_i\Big(b_{\sigma_i}(\theta)\Big)^2 + \sigma^2\sum_{i=1}^{k} q_i\Big(\partial_i b_{\sigma_i}(\theta)+1\Big)^2 \leq \sigma^2\sum_{i=1}^{k} q_i \tag{3.3}$$

αφού

$$\begin{aligned}
\mathbb{E}_{\theta,\sigma}\Big[\overline{q}(\phi(y)-\theta)\Big] &= \mathbb{E}_{\theta,\sigma}\Big[\sum_{i=1}^{k} q_i(\phi_i(y)-\theta_i)^2\Big] \\
&= \sum_{i=1}^{k} q_i\mathbb{E}_{\theta,\sigma}\Big[(\phi_i(y)-\theta_i)^2\Big] \\
&\geq \sum_{i=1}^{k} q_i\Big(b_{\sigma_i}(\theta)\Big)^2 + \sigma^2\sum_{i=1}^{k} q_i\Big(\partial_i b_{\sigma_i}(\theta)+1\Big)^2
\end{aligned}$$

και $tr(v_\Theta q) = \sum_{i=1}^{k} q_i$

Από την (3.3), με στοιχειώδεις υπολογισμούς, προκύπτει ότι:

$$\sigma^{-2}\sum_{i=1}^{k} q_i\Big(b_{\sigma_i}(\theta)\Big)^2 + 2\sum_{i=1}^{k} q_i\partial_i b_{\sigma_i}(\theta) + \sum_{i=1}^{k} q_i\Big(\partial_i b_{\sigma_i}(\theta)\Big)^2 \leq 0$$

Στη σχέση αυτή ο τελευταίος όρος είναι μη-αρνητικός, οπότε θα πρέπει:

$$\sigma^{-2}\sum_{i=1}^{k} q_i\Big(b_{\sigma_i}(\theta)\Big)^2 + 2\sum_{i=1}^{k} q_i\partial_i b_{\sigma_i}(\theta) \leq 0 \tag{3.4}$$

Από τη μορφή του εκτιμητή ϕ, που δίνεται από τον τύπο (3.1), λαμβάνουμε:

$$\begin{aligned}
b_\sigma(\theta) &= \mathbb{E}_{\theta,\sigma}\Big[\phi(y)-\theta\Big] \\
&= \mathbb{E}_{\theta,\sigma}\Big[\phi_0(y)-\theta-h\Big(\overline{v_\Theta^{-1}}(\phi_0(y)),s^2(y)\Big)c(\phi_0(y))\Big] \\
&= \underbrace{\mathbb{E}_{\theta,\sigma}\Big[\phi_0(y)-\theta\Big]}_{=0} - \mathbb{E}_{\theta,\sigma}\Big[h\Big(\overline{v_\Theta^{-1}}(\phi_0(y)),s^2(y)\Big)c(\phi_0(y))\Big] \\
&= -\mathbb{E}_{\theta,\sigma}\Big[h\Big(\overline{v_\Theta^{-1}}(\phi_0(y)),s^2(y)\Big)c(\phi_0(y))\Big] \\
&= -\int_{\Theta\times\Theta^\perp} h\Big(\overline{v_\Theta^{-1}}(x),\frac{\overline{v^{-1}}(z)}{n-k}\Big)c(x)N_\Theta(\theta,\sigma^2 v_\Theta;dx)N_{\Theta^\perp}(0,\sigma^2 v_{\Theta^\perp};dz) \\
&= -\int_\Theta c(x)\Big(\int_{\Theta^\perp} h\Big(\overline{v_\Theta^{-1}}(x),\frac{\overline{v^{-1}}(z)}{n-k}\Big)N_{\Theta^\perp}(0,\sigma^2 v_{\Theta^\perp};dz)\Big)N_\Theta(\theta,\sigma^2 v_\Theta;dx) \\
&= -\int_\Theta c(x)\Big(\int_{\Theta^\perp} h\Big(\overline{v_\Theta^{-1}}(x),\frac{\sigma^2}{n-k}\overline{v^{-1}}(z')\Big)N_{\Theta^\perp}(0,v_{\Theta^\perp};dz')\Big)N_\Theta(\theta,\sigma^2 v_\Theta;dx)
\end{aligned}$$

$$
\begin{aligned}
&= -\int_{\Theta} c(x)\Big(\int_{\mathbb{R}_+} h\Big(\overline{v_{\Theta}^{-1}}(x), \frac{\sigma^2}{n-k}u\Big)\mathcal{X}^2_{n-k}(du)\Big)N_{\Theta}(\theta, \sigma^2 v_{\Theta}; dx) \\
&= -\int_{\Theta} c(x)\Big(\int_{\mathbb{R}_+} h\Big(\overline{v_{\Theta}^{-1}}(x), u\Big)\Big(\frac{\sigma^2}{n-k}\mathcal{X}^2_{n-k}\Big)(du)\Big)N_{\Theta}(\theta, \sigma^2 v_{\Theta}; dx) \\
&= -\int_{\mathbb{R}_+}\Big(\int_{\Theta} c(x)h\Big(\overline{v_{\Theta}^{-1}}(x), u\Big)N_{\Theta}(\theta, \sigma^2 v_{\Theta}; dx)\Big)\Big(\frac{\sigma^2}{n-k}\mathcal{X}^2_{n-k}\Big)(du)
\end{aligned}
$$

Όμως, για κάθε u, εφαρμόζοντας το Λήμμα $5(iii)$ (βλ.Παράρτημα), έχουμε:

$$
\begin{aligned}
&\int_{\Theta} h\Big(\overline{v_{\Theta}^{-1}}(x), u\Big)c(x)N_{\Theta}(\theta, \sigma^2 v_{\Theta}; dx) \\
&= \int_{\Theta} h\Big(\sigma^2\overline{v_{\Theta}^{-1}}(x/\sigma), u\Big)\sigma c(x/\sigma)N_{\Theta}(\theta, \sigma^2 v_{\Theta}; dx) \\
&= \sigma\int_{\Theta} h\Big(\sigma^2\overline{v_{\Theta}^{-1}}(t), u\Big)c(t)N_{\Theta}\Big(\frac{\theta}{\sigma}, v_{\Theta}; dt\Big) \\
&= \sigma\int_{\mathbb{R}^k} h\Big(\sigma^2\|t\|^2, u\Big)c\,t\,N_k\Big(\frac{\theta}{\sigma}, I_k; dt\Big) \\
&= \sigma c\int_{\mathbb{R}^k} h\Big(\sigma^2\|t\|^2, u\Big)t\,N_k\Big(\frac{\theta}{\sigma}, I_k; dt\Big) \\
&= c\theta\int_{\mathbb{N}_0}\Big(\int_{\mathbb{R}_+} h(\sigma^2 v, u)\mathcal{X}^2_{k+2m+2}(dv)\Big)\pi\Big(\frac{\|\theta\|^2}{2\sigma^2}; dm\Big)
\end{aligned}
$$

Λοιπόν,

$$
\begin{aligned}
b_{\sigma}(\theta) &= -c\theta\int_{\mathbb{R}_+}\Big(\int_{\mathbb{N}_0}\Big(\int_{\mathbb{R}_+} h(\sigma^2 v, u)\mathcal{X}^2_{k+2m+2}(dv)\Big)\pi\Big(\frac{\|\theta\|^2}{2\sigma^2}; dm\Big)\Big)\Big(\frac{\sigma^2}{n-k}\mathcal{X}^2_{n-k}\Big)(du) \\
&= -c\theta\Psi_{\sigma}(\|\theta\|^2)
\end{aligned}
$$

όπου

$$
\Psi_{\sigma}(\|\theta\|^2) = e^{-\|\theta\|^2/2\sigma^2}\sum_{m=0}^{+\infty} H_m(\sigma^2)\frac{\|\theta\|^{2m}}{2^m\sigma^{2m}m!}
$$

με

$$
H_m(\sigma^2) = \int_{\mathbb{R}_+}\Big(\int_{\mathbb{R}_+} h(\sigma^2 v, u)\mathcal{X}^2_{k+2m+2}(dv)\Big)\Big(\frac{\sigma^2}{n-k}\mathcal{X}^2_{n-k}\Big)(du)
$$

Αυτό μας επιτρέπει να γράψουμε

$$
b_{\sigma}(\theta) = -\Psi_{\sigma}(t)c\theta
$$

όπου $t = \overline{v_{\Theta}^{-1}}(\theta)$ και η $\Psi_{\sigma}(t)$, όπως ορίστηκε παραπάνω, είναι διαφορίσιμη συνάρτηση, ως πρός t, από τον $\mathbb{R}_+$ στον $\mathbb{R}$.

Από την ανισότητα (3.4), λαμβάνουμε λοιπόν,

$$
-2\sum_{i=1}^{k} q_i\partial_i\Big[\Psi_{\sigma}(t)\sum_{j=1}^{k} c_{ij}\theta_j\Big] + \sigma^{-2}\sum_{i=1}^{k} q_i\Big[\Psi^2_{\sigma}(t)\Big(\sum_{j=1}^{k} c_{ij}\theta_j\Big)^2\Big] \leq 0
$$

δηλαδή,

$$-2\Psi_\sigma(t)\sum_{i=1}^{k} q_i c_{ii} - 2\Psi'_\sigma(t)\sum_{i=1}^{k} q_i \frac{\partial t}{\partial \theta_i}\sum_{j=1}^{k} c_{ij}\theta_j + \sigma^{-2}\sum_{i=1}^{k} q_i\Big[\Psi^2_\sigma(t)\Big(\sum_{j=1}^{k} c_{ij}\theta_j\Big)^2\Big] \leq 0$$

απ᾽ όπου προκύπτει:

$$2\Psi_\sigma(t)\sum_{i=1}^{k} q_i c_{ii} + 4\Psi'_\sigma(t)\sum_{i=1}^{k} q_i\theta_i\Big(\sum_{j=1}^{k} c_{ij}\theta_j\Big) - \sigma^{-2}\Psi^2_\sigma(t)\sum_{i=1}^{k} q_i\Big(\sum_{j=1}^{k} c_{ij}\theta_j\Big)^2 \geq 0$$

(αφού $\partial t/\partial\theta_i = 2\theta_i$).

Αυτή η συνθήκη μπορεί επιπλέον να συμπτυχθεί και να γραφεί στην μορφή:

$$2\Psi_\sigma(t)tr(v_\Theta qc) + 4\Psi'_\sigma(t)q\Big(\theta, c(\theta)\Big) - \sigma^{-2}\Psi^2_\sigma(t)\overline{q}\Big(c(\theta)\Big) \geq 0 \tag{3.5}$$

Θα αποδείξουμε ότι, αν η συνθήκη (3.2) δεν ικανοποιείται τότε η (3.5) έχει μοναδική λύση την $\Psi_\sigma = 0$.

Υποθέτουμε ότι υπάρχει $q \in Q$ τέτοιο ώστε $2pgvp(v_\Theta qc) \geq tr(v_\Theta qc)$.

1°) Έστω ότι υπάρχει t_0 τέτοιο ώστε $\Psi_\sigma(t_0) > 0$ και έστω f_1 ένα ιδιοδιάνυσμα του ενδομορφισμού $v_\Theta qc$ συνδεδεμένο με την ιδιοτιμή $pgvp(v_\Theta qc)$ και v_Θ^{-1}-ορθοκανονικό. Θεωρούμε τότε $\theta = t_0^{1/2} f_1$ (για το οποίο ισχύει $\overline{v_\Theta^{-1}}(\theta) = t_0$, αφού $\overline{v_\Theta^{-1}}(f_1) = 1$). Έχουμε, λοιπόν, ότι:

$$q\Big(\theta, c(\theta)\Big) = q\Big(t_0^{1/2} f_1, c(t_0^{1/2} f_1)\Big) = t_0 q\Big(f_1, c(f_1)\Big) = t_0\, {}^t f_1 qcf_1$$

και επειδή το f_1 είναι ιδιοδιάνυσμα που συνδέεται με την $pgvp(v_\Theta qc)$ έχουμε:

$$v_\Theta qcf_1 = pgvp(v_\Theta qc)f_1$$

δηλαδή,

$$qcf_1 = pgvp(v_\Theta qc)v_\Theta^{-1} f_1$$

οπότε,

$$q\Big(\theta, c(\theta)\Big) = t_0 pgvp(v_\Theta qc)^t f_1 v_\Theta^{-1} f_1 = t_0 pgvp(v_\Theta qc)\underbrace{\overline{v_\Theta^{-1}}(f_1)}_{=1} = t_0 pgvp(v_\Theta qc)$$

Επομένως, από την επιλογή αυτή των t_0 και θ, η (3.5) γίνεται:

$$2\Psi_\sigma(t_0)tr(v_\Theta qc) + 4\Psi'_\sigma(t_0)pgvp(v_\Theta qc)t_0 - \sigma^{-2}\Psi^2_\sigma(t_0)t_0\overline{q}(cf_1) \geq 0$$

Όμως, $\Psi_\sigma(t_0) > 0$ και $tr(v_\Theta qc) \leq 2pgvp(v_\Theta qc)$, άρα από την παραπάνω ανισότητα προκύπτει ότι:

$$4\Big(\Psi'_\sigma(t_0)t_0 + \Psi_\sigma(t_0)\Big)pgvp(v_\Theta qc) - \sigma^{-2}\Psi^2_\sigma(t_0)t_0\overline{q}(cf_1) \geq 0 \tag{3.6}$$

Επειδή, $-\sigma^{-2}\Psi^2_\sigma(t_0)t_0\overline{q}(cf_1) \leq 0$, ο πρώτος όρος της παραπάνω ανισότητας θα είναι μη αρνητικός, λοιπόν $\Psi'_\sigma(t_0)t_0 + \Psi_\sigma(t_0) \geq 0$ (επειδή $tr(v_\Theta qc) > 0$, εξ υποθέσεως) απ᾽ όπου προκύπτει ότι η $t \rightsquigarrow t\Psi_\sigma(t)$ είναι αύξουσα για κάθε $t \geq t_0$.

Πράγματι, αν υποθέσουμε ότι υπάρχει ένα διάστημα (t_1, t_2) με $t_0 < t_1 < t_2$ στο οποίο η $t\Psi_\sigma(t)$ είναι φθίνουσα για $t \in (t_1, t_2)$ και αύξουσα για $t \in (t_0, t_1)$ τότε θα έχουμε ότι:

$t_1\Psi_\sigma(t_1) \geq t_0\Psi_\sigma(t_0)$ απ᾽ όπου συμπεραίνουμε ότι $\Psi_\sigma(t_1) > 0$ και κάνοντας τα ίδια για το t_1 (όπως πρίν για το t_0) προκύπτει ο ίδιος τύπος:

$$4\Big(\Psi'_\sigma(t_1)t_1 + \Psi_\sigma(t_1)\Big)pgvp(v_\Theta qc) - \sigma^{-2}\Psi^2_\sigma(t_1)t_1\overline{q}(cf_1) \geq 0$$

απ᾽ όπου παίρνουμε $\Psi'_\sigma(t_1)t_1 + \Psi_\sigma(t_1) \geq 0$ και συνεπώς υπάρχει διάστημα (t_1, t_1') στο οποίο η $t \rightsquigarrow t\Psi_\sigma(t)$ είναι αύξουσα. Αυτό όμως είναι άτοπο και συνεπώς για κάθε $t \geq t_0$ η $t \rightsquigarrow t\Psi_\sigma(t)$ είναι αύξουσα.

Από το αποτέλεσμα αυτό και την ανισότητα (3.6) προκύπτει ότι, για κάθε $t \geq t_0$:

$$\frac{t\Psi'_\sigma(t) + \Psi_\sigma(t)}{\Psi^2_\sigma(t)} \geq \sigma^{-2}\overline{q}(cf_1)\frac{1}{4pgvp(v_\Theta qc)}\ t$$

Πολλαπλασιάζοντας και τα δύο μέλη με $1/t^2$ και ολοκληρώνοντας από t_0 έως t_1, λαμβάνουμε:

$$\Big(t_0\Psi_\sigma(t_0)\Big)^{-1} - \Big(t_1\Psi_\sigma(t_1)\Big)^{-1} \geq \frac{\sigma^{-2}\overline{q}(cf_1)}{4pgvp(v_\Theta qc)}\ \log\frac{t_1}{t_0}$$

και καταλήγουμε σε αντίφαση παίρνοντας $t_1 \to +\infty$ επειδή, το δεύτερο μέλος της ανισότητας τείνει στο $+\infty$ καθώς το $t_1 \to +\infty$, ενώ το πρώτο μέλος τείνει σε έναν σταθερό όρο, αφού

$$\lim_{t_1\to+\infty}\Big(t_1\Psi_\sigma(t_1)\Big)^{-1} = C\ (\geq 0)$$

2ο) Έστω ότι υπάρχει t_0 τέτοιο ώστε $\Psi_\sigma(t_0) < 0$, και έστω $(f_1,\ \ldots\ , f_k)$ μία βάση από ιδιοδιανύσματα του ενδομορφισμού $v_\Theta qc$, η οποία είναι v_Θ^{-1}-ορθοκανονική. Θεωρούμε τότε

$$\theta = \Big(\frac{t_0}{k}\Big)^{1/2}\sum_{i=1}^{k} f_i$$

(το οποίο επαληθεύει την: $\overline{v_\Theta^{-1}}(\theta) = t_0$). Έχουμε, λοιπόν, ότι:

$$q\Big(\theta, c(\theta)\Big) = q\bigg(\Big(\frac{t_0}{k}\Big)^{1/2}\sum_{i=1}^{k} f_i, \Big(\frac{t_0}{k}\Big)^{1/2}\sum_{j=1}^{k} c(f_j)\bigg) = \frac{t_0}{k}\sum_{i=1}^{k}\sum_{j=1}^{k} q\Big(f_i, c(f_j)\Big)$$

Όμως,

$$\begin{aligned} q\Big(f_i, c(f_j)\Big) &= {}^tf_iqcf_j = {}^tf_iv_\Theta^{-1}v_\Theta qcf_j = {}^tf_iv_\Theta^{-1}\lambda_jf_j \\ &= \lambda_j{}^tf_iv_\Theta^{-1}f_j = \lambda_jv_\Theta^{-1}(f_i, f_j) = \begin{cases} \lambda_j\ , & \text{αν } i = j \\ 0\ , & \text{αλλού} \end{cases} \end{aligned}$$

όπου λ_j, η ιδιοτιμή που συνδέεται με το ιδιοδιάνυσμα f_j. Επομένως,

$$q\Big(\theta, c(\theta)\Big) = \frac{t_0}{k}\sum_{j=1}^{k}\lambda_j = \frac{t_0}{k}tr(v_\Theta qc)$$

και συνεπώς η (3.5) γράφεται (για $t = t_0$ και το θεωρούμενο θ):

$$2\Psi_\sigma(t_0)tr(v_\Theta qc) + 4\Psi'_\sigma(t_0)\frac{t_0}{k}tr(v_\Theta qc) - \sigma^{-2}\Psi^2_\sigma(t_0)\frac{t_0}{k}\overline{q}\Big(\sum_{i=1}^{k} cf_i\Big) \geq 0$$

Η τελευταία αυτή ανισότητα (πολλαπλασιάζοντας με $k\Big(4tr(v_\Theta qc)\Big)^{-1} > 0$) γράφεται:

$$\frac{k}{2}\Psi_\sigma(t_0) + \Psi'_\sigma(t_0)t_0 - \alpha\Psi^2_\sigma(t_0)t_0 \geq 0 \tag{3.7}$$

με $\alpha = \dfrac{\sigma^{-2}}{4tr(v_\Theta qc)}\ \overline{q}\Big(\sum_{i=1}^{k} cf_i\Big) > 0$

Από την (3.7), επειδή $\Psi_\sigma(t_0) < 0$, συμπεραίνουμε ότι $\Psi'_\sigma(t_0) > 0$. Απ᾽ όπου προκύπτει ότι $\Psi_\sigma(t)$ είναι αύξουσα για κάθε $t \leq t_0$ (αποδεικνύεται όπως προηγουμένως για την $t \rightsquigarrow t\Psi_\sigma(t)$) και επομένως η (3.7) ισχύει για κάθε $t \leq t_0$.

Η (3.7) μπορεί να γραφεί ισοδύναμα (πολ/ζοντας και τα δύο μέλη με $t^{(k-2)/2}$) ως εξής:

$$\frac{k}{2}t^{(k-2)/2}\Psi_\sigma(t) + t^{k/2}\Psi'_\sigma(t) - \alpha\Psi^2_\sigma(t)t^{k/2} \geq 0$$

και διαιρώντας με $t^k\Psi^2_\sigma(t)$, έχουμε:

$$\frac{\frac{k}{2}t^{(k-2)/2}\Psi_\sigma(t) + t^{k/2}\Psi'_\sigma(t)}{t^k\Psi^2_\sigma(t)} - \alpha\ \frac{1}{t^{k/2}} \geq 0$$

απ᾽ όπου, ολοκληρώνοντας από t_1 έως t_0, παίρνουμε:

$$-\ \frac{1}{\Psi_\sigma(t_0)t_0^{k/2}}\ +\ \frac{1}{\Psi_\sigma(t_1)t_1^{k/2}}\ -\ \frac{2\alpha}{k-2}\Big(\frac{1}{t_1^{(k-2)/2}}\ -\ \frac{1}{t_0^{(k-2)/2}}\Big) \geq 0$$

Παίρνοντας το $t_1 \to 0$ καταλήγουμε σε αντίφαση επειδή:

$$\lim_{t_1\to 0}\frac{1}{\Psi_\sigma(t_1)t_1^{k/2}} \leq 0\ ,\ \text{ αφού } \forall t_1 \leq t_0,\ \ \frac{1}{\Psi_\sigma(t_1)t_1^{k/2}} < 0\ \ (\text{δηλ. προκύπτει } -\infty \geq 0\ ,\ \text{άτοπο})$$

Κατά συνέπεια, η μόνη λύση της (3.5) είναι η $\Psi_\sigma = 0$ και επομένως $b_\sigma(\theta) = 0$. Όμως τότε η σχέση που προέκυψε από την ανισότητα $Cramer - Rao$ γράφεται:

$$\mathbb{E}_{\theta,\sigma}\Big[\Big(\phi_i(y) - \theta_i\Big)^2\Big] \geq \sigma^2$$

και συνεπώς έχουμε:

$$\mathbb{E}_{\theta,\sigma}\Big[\overline{q}\Big(\phi(y) - \theta\Big)\Big] = \sum_{i=1}^{k} q_i\mathbb{E}_{\theta,\sigma}\Big[\Big(\phi_i(y) - \theta_i\Big)^2\Big] \geq \sigma^2\sum_{i=1}^{k} q_i = \sigma^2 tr(v_\Theta q)$$

Δηλαδή, δεν υπάρχει εκτιμητής της μορφής (3.1), εκτός από τον ϕ_0 που να κυριαρχεί τον ϕ_0 ομοιομόρφως.

□

Παρατήρηση

Επειδή ο εκτιμητής (3.1) είναι ειδική περίπτωση των εκτιμητών των προηγούμενων κεφαλαίων, εύλογα τίθεται το ερώτημα αν η χρήση μιας άλλης τετραγωνικής μορφής b, μέσα στην συνάρτηση συρρίκνωσης (δηλ. αν αυτή είχε τη μορφή $h\Big(\overline{b}(\phi_0(y)), s^2(y)\Big)$) θα επέτρεπε την ομοιόμορφη κυριαρχία επί του ϕ_0, ακόμα και αν η (3.2) δεν ικανοποιείται. Ιδιαίτερα, είδαμε στο Κεφ.2, ότι τέτοιοι εκτιμητές με συρρικνωτή "*matriciel*", για να είναι παραδεκτοί θα πρέπει $b = v_\Theta^{-1}c$. Παρατηρώντας ότι η ποσότητα $tr(v_\Theta qc) - 2pgvp(v_\Theta qc)$, η οποία εμφανίζεται και στο Κεφ.1, είναι ανεξάρτητη του b, φαίνεται δυνατή η επέκταση του παραπάνω θεωρήματος στην περίπτωση που b είναι μια οποιαδήποτε διγραμμική μορφή επί του Θ. Στην παρακάτω παράγραφο δίνουμε αυτήν την επέκταση για οποιαδήποτε b τέτοια ώστε $Kerb \subset Kerc$.

3.4 ΜΙΑ ΓΕΝΙΚΕΥΣΗ ΤΟΥ ΘΕΩΡΗΜΑΤΟΣ 3.3.1

Θεωρούμε την κλάση των εκτιμητών της μορφής:

$$\phi(y) = \Big(id_\Theta - h\Big(\overline{b}(\phi_0(y)), s^2(y)\Big)c\Big)\phi_0(y) \tag{3.8}$$

όπου b μια θετική συμμετρική διγραμμική μορφή επί του Θ, με $rg(b) \geq 3$ και c ένας ενδομορφισμός του Θ. Δεχόμαστε ότι $Kerb \subset Kerc$. Τότε:

Θεώρημα 3.4.1
Υπάρχει ένας εκτιμητής της μορφής (3.8) που κυριαρχεί τον ϕ_0 ομοιομόρφως στο (θ, σ^2) επί του Q αν και μόνον αν ικανοποιείται η (3.2).

Επειδή η απόδειξη αυτού του θεωρήματος είναι παρόμοια με εκείνη του Θεωρ.3.1.1, θα περιοριστούμε σε μία σκιαγράφηση αυτής, επιμένοντας μόνο στα σημεία στα οποία αυτή διαφοροποιείται.

Απόδειξη
Το ότι η συνθήκη (3.2) είναι ικανή εξασφαλίζεται από την Προτ.1.5.6, Κεφ.1, για $\mu_1 = 1$.

Έστω $(e_1, \ldots, e_k)$ μία βάση του Θ, v_Θ^{-1}-ορθοκανονική και q-ορθογώνια επί της οποίας η b διαγωνιοποιείται με διαγώνια στοιχεία 1 και 0 (δεχόμαστε ότι τα πρώτα $p(= rg(b))$ διαγώνια στοιχεία είναι 1) και ο c διαγωνιοποιείται με διαγώνια στοιχεία $c_i \geq 0$, $i = 1, ..., k$.

Τότε μέσα σ' αυτή τη βάση,

$$\theta = \sum_{i=1}^{k} \theta_i e_i \quad , \quad \overline{b}(\theta) = \sum_{i=1}^{p} \theta_i^2$$

και το διάνυσμα $\widetilde{b}_\sigma(\theta)$ της μεροληψίας γράφεται:

$$\mathbb{E}_{\theta,\sigma}\Big[\phi(y) - \theta\Big] = \widetilde{b}_\sigma(\theta) = \sum_{i=1}^{k} \widetilde{b}_{\sigma_i}(\theta) e_i$$

Σημειώνουμε με q_i, $i = 1, ..., k$ τις ιδιοτιμές του $v_\Theta q$.
Για να κυριαρχεί, λοιπόν, ο ϕ τον ϕ_0, θα πρέπει:

$$\sigma^{-2} \sum_{i=1}^{k} q_i \Big(\widetilde{b}_{\sigma_i}(\theta)\Big)^2 + 2 \sum_{i=1}^{k} q_i \partial_i \widetilde{b}_{\sigma_i}(\theta) \leq 0 \tag{3.9}$$

Λαμβάνοντας υπόψη τη μορφή του ϕ, η οποία δίδεται από την (3.8), έχουμε:

$$\begin{aligned}
\widetilde{b}_\sigma(\theta) &= -\mathbb{E}_{\theta,\sigma}\Big[h\Big(\overline{b}(\phi_0(y)), s^2(y)\Big)c(\phi_0(y))\Big] \\
&= -\int_E h\Big(\overline{b}(\phi_0(y)), s^2(y)\Big)c(\phi_0(y)) N_E(\theta, \sigma^2 v; dy) \\
&= -\int_\Theta c(x)\bigg(\int_{\Theta^\perp} h\Big(\overline{b}(x), \frac{1}{n-k}\overline{v^{-1}}(z)\Big) N_{\Theta^\perp}(0, \sigma^2 v_{\Theta^\perp}; dz)\bigg) N_\Theta(\theta, \sigma^2 v_\Theta; dx)
\end{aligned}$$

$$= \ -\int_{\Theta} c(x)\Big(\int_{\Theta^\perp} h\Big(\overline{b}(x), \frac{\sigma^2}{n-k}\overline{v^{-1}}(z)\Big) N_{\Theta^\perp}(0, v_{\Theta^\perp}; dz)\Big) N_\Theta(\theta, \sigma^2 v_\Theta; dx)$$

$$= \ -\int_{\Theta} c(x)\Big(\int_{\mathbb{R}_+} h\Big(\overline{b}(x), \frac{\sigma^2}{n-k}u\Big)\mathcal{X}^2_{n-k}(du)\Big) N_\Theta(\theta, \sigma^2 v_\Theta; dx)$$

$$= \ -\int_{\mathbb{R}_+}\Big(\int_{\Theta} h\Big(\overline{b}(x), u\Big) c(x) N_\Theta(\theta, \sigma^2 v_\Theta; dx)\Big)\Big(\frac{\sigma^2}{n-k}\mathcal{X}^2_{n-k}\Big)(du)$$

Όμως,

$$\begin{aligned}\int_{\Theta} h\Big(\overline{b}(x), u\Big) c(x) N_\Theta(\theta, \sigma^2 v_\Theta; dx) &= \sigma\int_{\Theta} h\Big(\sigma^2\overline{b}(t), u\Big)c(t)N_\Theta\Big(\frac{\theta}{\sigma}, v_\Theta; dt\Big) \\ &= \sigma\int_{\mathbb{R}^k} h\Big(\sigma^2\overline{b}(t), u\Big)ctN_k\Big(\frac{\theta}{\sigma}, I_k; dt\Big) \qquad (3.10)\end{aligned}$$

Παρατηρούμε ότι: $\forall i \in \{1, 2, ..., k\}$,

$$\begin{aligned}&\frac{\partial}{\partial\theta_i}\int_{\mathbb{R}^k} h\Big(\sigma^2\overline{b}(t), u\Big)N_k\Big(\frac{\theta}{\sigma}, I_k; dt\Big) \\ &= \frac{\partial}{\partial\theta_i}\int_{\mathbb{R}^k} h\Big(\sigma^2\overline{b}(t), u\Big)(2\pi)^{-k/2}exp\Big\{-\frac{1}{2}\sum_{i=1}^{k}\Big(t_i - \frac{\theta_i}{\sigma}\Big)^2\Big\}dt_1...dt_k \\ &= \frac{1}{\sigma}\int_{\mathbb{R}^k} h\Big(\sigma^2\overline{b}(t), u\Big)\Big(t_i - \frac{\theta_i}{\sigma}\Big)N_k\Big(\frac{\theta}{\sigma}, I_k; dt\Big) \qquad (3.11)\end{aligned}$$

Επίσης, επειδή ο πίνακας b ικανοποιεί $b \cdot b = b$, έχουμε ότι, η εικόνα του νόμου $N_k(\theta/\sigma, I_k)$ υπό την τετραγωνική μορφή $\overline{b}$ είναι ο νόμος $\mathcal{X}^2_p(\delta)$ με $p = rg(b)$ και $\delta = \frac{1}{\sigma^2}\overline{b}(\theta)$(βλ.Λήμμα 4, Παράρτημα).

Λοιπόν, μπορούμε να γράψουμε:

$$\begin{aligned}\int_{\mathbb{R}^k} h\Big(\sigma^2\overline{b}(t), u\Big)N_k\Big(\frac{\theta}{\sigma}, I_k; dt\Big) &= \int_{\mathbb{R}_+} h(\sigma^2 s, u)\mathcal{X}^2_p\Big(\frac{\overline{b}(\theta)}{\sigma^2}; ds\Big) \\ &= \int_{\mathbb{N}_0}\Big(\int_{\mathbb{R}_+} h(\sigma^2 s, u)\mathcal{X}^2_{p+2m}(ds)\Big)\pi\Big(\frac{\overline{b}(\theta)}{2\sigma^2}; dm\Big) \qquad (3.12)\end{aligned}$$

Ονομάζοντας

$$f_\sigma(m, u) = \int_{\mathbb{R}_+} h(\sigma^2 s, u)\mathcal{X}^2_{p+2m}(ds)$$

η (3.12), δίνει : $\forall i \in \{1, 2, ..., p\}$,

$$\begin{aligned}&\frac{\partial}{\partial\theta_i}\int_{\mathbb{R}^k} h\Big(\sigma^2\overline{b}(t), u\Big)N_k\Big(\frac{\theta}{\sigma}, I_k; dt\Big) \\ &= \frac{\partial}{\partial\theta_i}\sum_{m=0}^{+\infty} f_\sigma(m, u)exp\Big\{-\frac{1}{2\sigma^2}\sum_{i=1}^{p}\theta_i^2\Big\}\Big(\frac{1}{2\sigma^2}\sum_{i=1}^{p}\theta_i^2\Big)^m\frac{1}{m!}\end{aligned}$$

$$= \sum_{m=0}^{+\infty} \frac{f_\sigma(m,u)}{m!} \Big[-\frac{1}{\sigma^2}\theta_i exp\Big\{-\frac{1}{2\sigma^2}\sum_{i=1}^{p}\theta_i^2\Big\}\Big(\frac{1}{2\sigma^2}\sum_{i=1}^{p}\theta_i^2\Big)^m$$

$$+ exp\Big\{-\frac{1}{2\sigma^2}\sum_{i=1}^{p}\theta_i^2\Big\} m \Big(\frac{1}{2\sigma^2}\sum_{i=1}^{p}\theta_i^2\Big)^{m-1}\frac{\theta_i}{\sigma^2}\Big]$$

$$= \sum_{m=0}^{+\infty} \frac{f_\sigma(m,u)}{m!}\frac{\theta_i}{\sigma^2} exp\Big\{-\frac{1}{2\sigma^2}\sum_{i=1}^{p}\theta_i^2\Big\}\Big(\frac{1}{2\sigma^2}\sum_{i=1}^{p}\theta_i^2\Big)^{m-1}\Big(m - \frac{1}{2\sigma^2}\sum_{i=1}^{p}\theta_i^2\Big)$$

$$= \frac{\theta_i}{\sigma^2}\sum_{m=0}^{+\infty}\frac{f_\sigma(m,u)}{m!} exp\Big\{-\frac{\overline{b}(\theta)}{2\sigma^2}\Big\}\Big(\frac{\overline{b}(\theta)}{2\sigma^2}\Big)^{m-1}\Big(m - \frac{\overline{b}(\theta)}{2\sigma^2}\Big)$$

$$= \frac{\theta_i}{\overline{b}(\theta)}\sum_{m=0}^{+\infty}\frac{f_\sigma(m,u)}{m!} exp\Big\{-\frac{\overline{b}(\theta)}{2\sigma^2}\Big\}\Big(\frac{\overline{b}(\theta)}{2\sigma^2}\Big)^{m}\Big(2m - \frac{\overline{b}(\theta)}{\sigma^2}\Big)$$

$$= \frac{\theta_i}{\overline{b}(\theta)}\int_{\mathbb{N}_0} f_\sigma(m,u)\Big(2m - \frac{\overline{b}(\theta)}{\sigma^2}\Big)\pi\Big(\frac{\overline{b}(\theta)}{2\sigma^2}; dm\Big) \tag{3.13}$$

Εξισώνοντας τις (3.11) και (3.13) προκύπτει ότι:

$$\frac{1}{\sigma}\int_{\mathbb{R}^k} h\Big(\sigma^2\overline{b}(t), u\Big)\Big(t_i - \frac{\theta_i}{\sigma}\Big)N_k\Big(\frac{\theta}{\sigma}, I_k; dt\Big) = \frac{\theta_i}{\overline{b}(\theta)}\int_{\mathbb{N}_0} f_\sigma(m,u)\Big(2m - \frac{\overline{b}(\theta)}{\sigma^2}\Big)\pi\Big(\frac{\overline{b}(\theta)}{2\sigma^2}; dm\Big)$$

από το οποίο, λόγω της (3.12), λαμβάνουμε: $\forall i \in \{1, 2, ..., p\}$,

$$\frac{1}{\sigma}\int_{\mathbb{R}^k} h\Big(\sigma^2\overline{b}(t), u\Big)t_i N_k\Big(\frac{\theta}{\sigma}, I_k; dt\Big) = \frac{\theta_i}{\overline{b}(\theta)}\int_{\mathbb{N}_0} f_\sigma(m,u)2m\pi\Big(\frac{\overline{b}(\theta)}{2\sigma^2}; dm\Big)$$

Επειδή $rg(c) \leq rg(b)$, p το πολύ στοιχεία c_i είναι μη-μηδενικά και επομένως η (3.10) γίνεται:

$$\int_\Theta h\Big(\overline{b}(x), u\Big)c(x)N_\Theta(\theta, \sigma^2 v_\Theta; dx) = \sigma^2\frac{c\theta}{\overline{b}(\theta)}\int_{\mathbb{N}_0} f_\sigma(m,u)2m\pi\Big(\frac{\overline{b}(\theta)}{2\sigma^2}; dm\Big)$$

Έτσι, το διάνυσμα της μεροληψίας $\widetilde{b}_\sigma(\theta)$, γράφεται:

$$\begin{aligned}\widetilde{b}_\sigma(\theta) &= -c\theta\int_{\mathbb{R}_+}\frac{\sigma^2}{\overline{b}(\theta)}\Big(\int_{\mathbb{N}_0} f_\sigma(m,u)2m\pi\Big(\frac{\overline{b}(\theta)}{2\sigma^2}; dm\Big)\Big)\Big(\frac{\sigma^2}{n-k}\mathcal{X}^2_{n-k}\Big)(du)\\ &= -\Psi_\sigma(t)c\theta \quad \text{με} \;\; t = \overline{b}(\theta)\end{aligned} \tag{3.14}$$

όπου

$$\Psi_\sigma(t) = e^{-t^2/2\sigma^2}\sum_{m=0}^{+\infty} H_m(\sigma^2)\frac{t^m}{2^m\sigma^{2m}m!} \;\; ,$$

με

$$H_m(\sigma^2) = \int_{\mathbb{R}_+}\Big(\int_{\mathbb{R}_+} h(\sigma^2 s, u)\mathcal{X}^2_{p+2m+2}(ds)\Big)\Big(\frac{\sigma^2}{n-k}\mathcal{X}^2_{n-k}\Big)(du) \;\; ,$$

διαφορίσιμη, ως πρός t, συνάρτηση από τον $\mathbb{R}_+$ στον $\mathbb{R}$.

Δεχόμαστε ότι τα πρώτα p -το πολύ στοιχεία του πίνακα του ενδομορφισμού c είναι μη-μηδενικά. Τότε η (3.9), λόγω της (3.14), γράφεται:

$$\sigma^{-2}\sum_{i=1}^{k} q_i \Psi_\sigma^2(t)\Big(\sum_{j=1}^{k} c_{ij}\theta_j\Big)^2 + 2\sum_{i=1}^{k} q_i \frac{\partial}{\partial\theta_i}\Big(-\Psi_\sigma(t)\sum_{j=1}^{k} c_{ij}\theta_j\Big) \le 0$$

ισοδύναμα

$$\sigma^{-2}\Psi_\sigma^2(t)\sum_{i=1}^{p} q_i (c_i\theta_i)^2 + 2\sum_{i=1}^{p} q_i \frac{\partial}{\partial\theta_i}\Big(-\Psi_\sigma(t)c_i\theta_i\Big) \le 0$$

δηλαδή

$$2\Psi_\sigma(t)\sum_{i=1}^{p} q_i c_i + 4\Psi'_\sigma(t)\sum_{i=1}^{p} q_i c_i\theta_i^2 - \sigma^{-2}\Psi_\sigma^2(t)\sum_{i=1}^{p} q_i c_i^2\theta_i^2 \ge 0$$

ή συνοπτικότερα

$$2\Psi_\sigma(t)tr(v_\Theta qc) + 4\Psi'_\sigma(t)bv_\Theta q\Big(\theta, c(\theta)\Big) - \sigma^{-2}\Psi_\sigma^2(t)\overline{q}\Big(c(\theta)\Big) \ge 0 \qquad (3.15)$$

παρατηρώντας ότι $q(\theta, c(\theta)) = bv_\Theta q(\theta, c(\theta))$.

Αποδεικνύεται ομοίως με το Θεώρ.3.1.3, ότι αν η συνθήκη (3.2) δεν ικανοποιείται η (3.15) έχει μοναδική λύση την $\Psi_\sigma = 0$. Υποθέτουμε λοιπόν ότι υπάρχει $q \in Q$, τέτοιο ώστε:

$$2pgvp(v_\Theta qc) \ge tr(v_\Theta qc)$$

1ο) Έστω ότι υπάρχει t_0 τέτοιο ώστε $\Psi_\sigma(t_0) > 0$ και έστω f_1 ένα ιδιοδιάνυσμα του ενδομορφισμού $v_\Theta qc$ συνδεδεμένο με την ιδιοτιμή $pgvp(v_\Theta qc)$ και b-ορθοκανονικό. Θεωρούμε τότε $\theta = t_0^{1/2} f_1$ (για το οποίο ισχύει, $\overline{b}(\theta) = t_0$). Έχουμε, λοιπόν, ότι:

$$\begin{aligned} bv_\Theta q\Big(\theta, c(\theta)\Big) &= bv_\Theta q\Big(t_0^{1/2} f_1, c(t_0^{1/2} f_1)\Big) \\ &= t_0 bv_\Theta q\Big(f_1, c(f_1)\Big) \\ &= t_0\,{}^t f_1 b(v_\Theta qcf_1) \\ &= t_0 pgvp(v_\Theta qc)\,{}^t f_1 b f_1 \\ &= t_0 pgvp(v_\Theta qc) \end{aligned}$$

Επομένως για την επιλογή αυτή των t_0 και θ, η (3.15) καταλήγει στην (3.6) του Θεωρ.3.1.1, η οποία οδηγεί σε αντίφαση.

2ο) Έστω ότι υπάρχει t_0 τέτοιο ώστε $\Psi_\sigma(t_0) < 0$, και έστω $(f_1, \ \ldots\ , f_k)$ μία βάση από ιδιοδιανύσματα του ενδομορφισμού $v_\Theta qc$, η οποία είναι b-ορθοκανονική. Θεωρούμε τότε

$$\theta = \Big(\frac{t_0}{k}\Big)^{1/2}\sum_{i=1}^{k} f_i$$

(το οποίο επαληθεύει: $\overline{b}(\theta) = \frac{t_0}{k}\overline{b}\Big(\sum_{i=1}^{k} f_i\Big) = \frac{t_0}{k}\sum_{i=1}^{k}\sum_{j=1}^{k} b(f_i, f_j) = t_0$).

Τότε, λοιπόν, έχουμε:

$$\begin{aligned}
bv_\Theta q\Big(\theta, c(\theta)\Big) &= bv_\Theta q\Bigg(\Big(\frac{t_0}{k}\Big)^{1/2}\sum_{i=1}^{k} f_i, \Big(\frac{t_0}{k}\Big)^{1/2} c\Big(\sum_{j=1}^{k} f_j\Big)\Bigg) \\
&= \frac{t_0}{k} bv_\Theta q\Big(\sum_{i=1}^{k} f_i, \sum_{j=1}^{k} c(f_j)\Big) \\
&= \frac{t_0}{k}\sum_{i=1}^{k}\sum_{j=1}^{k} bv_\Theta q\Big(f_i, c(f_j)\Big)
\end{aligned}$$

Όμως,

$$\begin{aligned}
bv_\Theta q\Big(f_i, c(f_j)\Big) &= {}^t f_i bv_\Theta q c f_j \\
&= \lambda_j f_i b f_j = \begin{cases} \lambda_j \ , & \text{αν } i = j \\ 0 \ , & \text{αλλού} \end{cases}
\end{aligned}$$

όπου λ_j η ιδιοτιμή που αντιστοιχεί στο ιδιοδιάνυσμα f_j. Επομένως,

$$bv_\Theta q\Big(\theta, c(\theta)\Big) = \frac{t_0}{k}\sum_{j=1}^{k}\lambda_j = \frac{t_0}{k} tr(v_\Theta q c)$$

Για την επιλογή λοιπόν αυτή των t_0 και θ, η (3.15) καταλήγει στην (3.7) του Θεωρ.3.1.1, η οποία οδηγεί σε αντίφαση.

$\square$

Κεφάλαιο 4

ΣΥΓΚΡΙΣΕΙΣ ΜΕΤΑΞΥ ΤΩΝ ΕΚΤΙΜΗΤΩΝ ΜΕ ΣΥΡΡΙΚΝΩΤΗ

4.1 ΓΕΝΙΚΟΤΗΤΕΣ - OPTIMALITÉ

4.1.1

Στο κεφάλαιο αυτό περιοριζόμαστε στους εκτιμητές των προηγούμενων κεφαλαίων, δηλαδή σ' αυτούς των *Fraisse*, *Robert* και *Roy*[1987] και *Cellier*, *Fourdrinier* και *Robert*[1989], οι οποίοι γενικεύουν τα αποτελέσματα των *Judge* και *Bock*[1978] αντίστοιχα στις περιπτώσεις όπου οι διανυσματικοί τελεστές που υπεισέρχονται στον εκτιμητή είναι οποιοιδήποτε και όπου η συνάρτηση συρρίκνωσης δεν είναι αναγκαστικά συνεχής.

Λίγοι σχετικά συγγραφείς, έχουν αποδείξει προοδευτικά και εκτενώς συνθήκες κυριαρχίας στο εσωτερικό τέτοιων κλάσεων. Εκτός από τα ακριβή αποτελέσματα (παραδοχή του προτεινόμενου εκτιμητή, κυριαρχία του εκτιμητή των *James-Stein*(βλ.*Tze* και *Wen*[1982])), μπορούμε να παραθέσουμε την αναγκαία συνθήκη παραδοχής αποδεδειγμένη από τους *Berger* και *Srinivasan*[1978] η οποία γενικεύεται στο Κεφ.2 και την απόδειξη για την κυριαρχία του "*positive rule*" εκτιμητή την οποία μπορούμε να βρούμε υπό μία μορφή σχετικά γενική στους *Judge* και *Bock*[1978, σελ.238-240].

Προτού εξετάσουμε τα αποτελέσματα επί της επίδοσης ενός εκτιμητή με συρρικνωτή τροποποιώντας ή κόβοντας τη συνάρτηση συρρικνώσεως του, θα εξετάσουμε υπό ποιές συνθήκες υπάρχει βέλτιστη(*optimal*) τροποποίηση κλίμακας του όρου συρρικνώσεως. Το μοντέλο το οποίο θεωρούμε είναι εκείνο της εκτίμησης του μέσου ενός κανονικού διανύσματος με διασπορά γνωστή εκτός από έναν πολλαπλασιαστικό παράγοντα. Οι εκτιμητές με συρρικνωτή τους οποίους θεωρούμε είναι της μορφής:

$$\phi(y) = \Big(id_\Theta - h\Big(\overline{b}(\phi_0(y)), 1/s^2(y)\Big)c\Big)\phi_0(y)$$

με τις υποθέσεις και τις έννοιες των *Cellier*, *Fourdrinier* και *Robert*[1989]. Ιδιαιτέρως, υποθέτουμε ότι b και c διαγωνιοποιούνται ταυτόχρονα σε μια βάση v_Θ^{-1}-ορθοκανονική και q-ορθογώνια. Η συνάρτηση συρρικνώσεως h θα θεωρείται δεδομένη.

4.1.2 Βελτιστοποίηση με την έννοια των James – Stein

Κατά έναν κλασσικό τρόπο, όπως εκείνον των *James – Stein*[1961] οι οποίοι αποδεικνύουν ότι ο εκτιμητής:

$$\phi_{JS}(y) = \Big(1 - \rho^* \frac{1}{tu}\Big)\phi_0(y)$$

(όπου $\rho^* = \frac{(k-2)(n-k)}{n-k+2}$, $t = \overline{b}(\phi_0(y))$, $u = 1/s^2(y)$ και $q = b = v_\Theta^{-1}$) είναι βέλτιστος, με τη συνήθη έννοια του τετραγωνικού σφάλματος, μεταξύ των εκτιμητών:

$$\phi_\rho(y) = \Big(1 - \rho\frac{1}{tu}\Big)\phi_0(y)$$

αποδεικνύεται ότι,για την κλάση των θεωρούμενων εκτιμητών, δεν υπάρχει βέλτιστος εκτιμητής, με την έννοια των $James - Stein$, όπως δείχνει το ακόλουθο αποτέλεσμα.

Πρόταση 4.1.1

Μεταξύ των εκτιμητών

$$\phi_\lambda(y) = \Big(id_\Theta - \lambda h(t,u)c\Big)\phi_0(y)\ , \quad \lambda \in \mathbb{R} \tag{4.1}$$

ο εκτιμητής ελαχίστου κινδύνου στο (θ, σ) συνδέεται με το:

$$\lambda(\theta,\sigma) = \frac{\mathbb{E}_{\theta,\sigma}\Big[h(t,u)\sum_{i=1}^k q_i c_i y_i (y_i - \theta_i)\Big]}{\mathbb{E}_{\theta,\sigma}\Big[h^2(t,u)\sum_{i=1}^k q_i c_i^2 y_i^2\Big]}$$

Απόδειξη

Για τον κίνδυνο που συνδέεται με την τετραγωνική απώλεια $\sigma^{-2}\overline{q}(\widehat{\theta} - \theta)$, έχουμε:

$$\begin{aligned}
& R_{\phi_0}(\theta,\sigma) - R_{\phi_\lambda}(\theta,\sigma) \\
&= \frac{2}{\sigma^2}\ \mathbb{E}_{\theta,\sigma}\Big[q\Big(\phi_0(y) - \theta, \lambda h(t,u)c(\phi_0(y))\Big)\Big] - \frac{1}{\sigma^2}\ \mathbb{E}_{\theta,\sigma}\Big[\overline{q}\Big(\lambda h(t,u)c(\phi_0(y))\Big)\Big] \\
&= \frac{2}{\sigma^2}\ \mathbb{E}_{\theta,\sigma}\Big[\lambda h(t,u)\sum_{i=1}^k q_i c_i (y_i - \theta_i) y_i\Big] - \frac{1}{\sigma^2}\ \mathbb{E}_{\theta,\sigma}\Big[\lambda^2 h^2(t,u)\sum_{i=1}^k q_i c_i^2 y_i^2\Big] \\
&= \lambda\Big[\frac{2}{\sigma^2}\ \mathbb{E}_{\theta,\sigma}\Big[h(t,u)\sum_{i=1}^k q_i c_i y_i (y_i - \theta_i)\Big] - \frac{\lambda}{\sigma^2}\ \mathbb{E}_{\theta,\sigma}\Big[h^2(t,u)\sum_{i=1}^k q_i c_i^2 y_i^2\Big]\Big]
\end{aligned}$$

επειδή, ως προς τη θεωρούμενη βάση (βλ. τέλος της §4.1.1),

$$\phi_0(y) = \sum_{i=1}^k y_i e_i$$

και

$$\begin{aligned}
q\Big(\phi_0(y) - \theta, c(\phi_0(y))\Big) &= q\Big(\sum_{i=1}^k (y_i - \theta_i)e_i, c\Big(\sum_{i=1}^k y_i e_i\Big)\Big) \\
&= \sum_{i=1}^k q\Big((y_i - \theta_i)e_i, \sum_{j=1}^k y_j c(e_j)\Big) \\
&= \sum_{i=1}^k (y_i - \theta_i) q\Big(e_i, \sum_{j=1}^k y_j c_j e_j\Big)
\end{aligned}$$

$$= \sum_{i=1}^{k}\sum_{j=1}^{k}(y_i-\theta_i)y_jc_jq(e_i,e_j)$$

$$= \sum_{i=1}^{k} q_ic_iy_i(y_i-\theta_i)$$

Όμως από την προηγούμενη σχέση της διαφοράς των κινδύνων είναι φανερό οτι η ποσότητα αυτή μεγιστοποιείται όταν το λ πάρει την τιμή $\lambda(\theta,\sigma)$.

□

Το $\lambda(\theta,\sigma)$ εξαρτόμενο γενικά από το (θ,σ), σημαίνει ότι δεν υπάρχει για το πλείστον των συναρτήσεων συρρικνώσεως, εκτιμητής της κλάσης (4.1) ο οποίος να είναι ελαχίστου κινδύνου ομοιομόρφως στο (θ,σ). Ξαναβρίσκουμε λοιπόν, για $h(t,u)=1/tu$, τη μερική περίπτωση $1.4.2(B)$ του Κεφ.1 η οποία παρέχει μια κλάση εκτιμητών με συρρικνωτή οι οποίοι δεν είναι συγκρίσιμοι (Προτ.1.4.5, Κεφ.1).

4.2 ΕΝΙΣΧΥΣΗ ΤΗΣ ΣΥΡΡΙΚΝΩΣΗΣ

4.2.1 Γενική Περίπτωση

Δεδομένης μιας συνάρτησης συρρικνώσεως h, θα εξετάσουμε υπό ποιά συνθήκη μία τροποποίηση του ενδομορφισμού συρρικνώσεως επιτρέπει, μία ομοιόμορφη ως προς (θ,σ), ελλάτωση του κινδύνου του συνδεδεμένου με αυτόν εκτιμητή. Θα λάβουμε υπόψη μόνο την 'ένίσχυση'' της συρρίκνωσης, δηλαδή μία αύξηση των ιδιοτιμών του ενδομορφισμού συρρικνώσεως, θεωρώντας ότι, με δεδομένη τη συνάρτηση συρρικνώσεως, μπορούμε πάντοτε να ξεκινήσουμε από έναν ενδομορφισμό με ιδιοτιμές ''μικρές'' και τις οποίες αυξάνουμε όσο ο περιορισμός που εισάγεται στην παρακάτω πρόταση ικανοποιείται.

Θεωρούμε τους εκτιμητές της μορφής:

$$\phi(y) = \Big(id_\Theta - h(t,u)c\Big)\phi_0(y)\ ,\ \ \text{με}\ \ t=\bar{b}\Big(\phi_0(y)\Big),\ u=\frac{1}{s^2(y)} \tag{4.2}$$

όπου b θετική, συμμετρική, διγραμμική μορφή επί του Θ,
c ενδομορφισμός του Θ,
$h:\mathbb{R}_+^2 \to \mathbb{R}_+$ μετρήσιμη συνάρτηση

και υποθέτουμε, συμφωνα με τους *Cellier*, *Fourdrinier* και *Robert*[1989], ότι τα b, c και q ικανοποιούν τις παρακάτω υποθέσεις:

$(\tilde{P}1)$ Υπάρχει μια βάση $(e_1,...,e_n)$ του Ε, v^{-1}-ορθοκανονική, της οποίας τα k πρώτα στοιχεία να αποτελούν βάση του Θ, q-ορθογώνια, μέσα στην οποία b και c διαγωνιοποιούνται.

$(\tilde{P}2)$ Όλες οι ιδιοτιμές του c είναι μη αρνητικές.

$(\tilde{P}3)$ Η $rg(b)\geq 3$, το οποίο υποδηλώνει ότι $k\geq 3$, και $Kerb\subset Kerc$, το οποίο υποδηλώνει ότι ο περιορισμός του cb^- στην Imb ορίζεται μονοσήμαντα για κάθε γενικευμένο αντίστροφο b^- του b.

$(\tilde{P}4)$ ${}^tcqcb^-$ δεν είναι ταυτοτικά μηδέν.

Επιπλέον υποθέτουμε ότι η h επαληθεύει τις υποθέσεις του "ελέγχου":

(a) Υπάρχει λ_1 τέτοιο ώστε η $h(\cdot, u)$ να είναι ελεγχόμενη σε ύψος λ_1, για κάθε u.

(b) Υπάρχει λ_2 τέτοιο ώστε η $h(t, \cdot)$ να είναι ελεγχόμενη σε ύψος λ_2, για κάθε t.

Τότε (βλ. βασικό Θεώρημα στο [16]) έχουμε:

Λήμμα

Υπο τις υποθέσεις $(\widetilde{P}1)$,$(\widetilde{P}2)$,$(\widetilde{P}3)$ *και* $(\widetilde{P}4)$, *μια ικανή συνθήκη για να κυριαρχεί ο εκτιμητής* ϕ *που ορίζεται από την (4.2) τον εκτιμητή ελαχίστων τετραγώνων* ϕ_0 *είναι η* h *να ικανοποιεί τις συνθήκες ελέγχου* (a), (b) *και η απεικόνιση*

$$(t, u) \rightsquigarrow tuh(t, u)$$

να είναι ομοιομόρφως φραγμένη από την ποσότητα:

$$\beta = 2\, \frac{tr(v_\Theta qc) - 2\lambda_1 pgvp(v_\Theta qc)}{pgvp(\,{}^t cqcb^-)} \;\; \frac{n-k}{n-k+2\lambda_2}$$

Θεωρώντας έναν εκτιμητή της μορφής (4.2) και $\epsilon = (\epsilon_1, \ \ldots \ , \epsilon_k) \in (\mathbb{R}_+)^k$ εισάγουμε τον *διαταραγμένο* (*perturbé*) εκτιμητή:

$$\phi_\epsilon(y) = \Big(id_\Theta - h(t, u) c_\epsilon \Big) \phi_0(y)$$

όπου ο ενδομορφισμός c_ϵ δέχεται τον πίνακα:

$$\begin{pmatrix} c_1 + \epsilon_1 & & 0 \\ & \ddots & \\ 0 & & c_k + \epsilon_k \end{pmatrix}$$

μέσα στην βάση ιδιοδιανυσμάτων του c η οποία είναι v_Θ^{-1}-ορθοκανονική και q- ορθογώνια. Μέσα σ'αυτή τη βάση, για $1 \le i \le k$, το θ_i εκτιμάται από το $y_i - h(t, u)(c_i + \epsilon_i) y_i$, γι'αυτόν τον διαταραγμένο εκτιμητή.

Λαμβάνουμε, λοιπόν, το ακόλουθο αποτέλεσμα:

Πρόταση 4.2.2

Ο ϕ_ϵ *κυριαρχεί τον* ϕ *ομοιομόρφως στο* (θ, σ), *αν*

$$\sup_{t,u} tuh(t, u) \le 2\, \frac{\sum_{i=1}^k q_i \epsilon_i - 2\lambda_1 \max(q_i \epsilon_i)}{\max\{q_i b_i^{-1} \epsilon_i (2c_i + \epsilon_i)\}} \; \frac{n-k}{n-k+2\lambda_2}$$

Απόδειξη

Η διαφορά των κινδύνων μεταξύ των δύο εκτιμητών, γράφεται:

$$\begin{aligned} & R_\phi(\theta, \sigma) - R_{\phi_\epsilon}(\theta, \sigma) \\ &= -\frac{2}{\sigma^2} \mathbb{E}_{\theta,\sigma} \Big[h(t, u) \sum_{i=1}^k q_i c_i (y_i - \theta_i) y_i \Big] + \frac{1}{\sigma^2} \mathbb{E}_{\theta,\sigma} \Big[h^2(t, u) \sum_{i=1}^k q_i c_i^2 y_i^2 \Big] \end{aligned}$$

$$+\frac{2}{\sigma^2}\mathbb{E}_{\theta,\sigma}\Big[h(t,u)\sum_{i=1}^{k} q_i(c_i+\epsilon_i)(y_i-\theta_i)y_i\Big] - \frac{1}{\sigma^2}\mathbb{E}_{\theta,\sigma}\Big[h^2(t,u)\sum_{i=1}^{k} q_i(c_i+\epsilon_i)^2 y_i^2\Big]$$

$$= \frac{2}{\sigma^2}\mathbb{E}_{\theta,\sigma}\Big[h(t,u)\sum_{i=1}^{k} q_i(c_i+\epsilon_i-c_i)(y_i-\theta_i)y_i\Big] - \frac{1}{\sigma^2}\mathbb{E}_{\theta,\sigma}\Big[h^2(t,u)\sum_{i=1}^{k} q_i\Big((c_i+\epsilon_i)^2-c_i^2\Big)y_i^2\Big]$$

$$= \frac{2}{\sigma^2}\mathbb{E}_{\theta,\sigma}\Big[h(t,u)\sum_{i=1}^{k} q_i\epsilon_i y_i(y_i-\theta_i)\Big] - \frac{1}{\sigma^2}\mathbb{E}_{\theta,\sigma}\Big[h^2(t,u)\sum_{i=1}^{k} q_i\epsilon_i(2c_i+\epsilon_i)y_i^2\Big]$$

Θέτουμε:

$$\widetilde{q_i} = q_i\ \frac{\epsilon_i}{2c_i+\epsilon_i}\ ,\ \ \widetilde{c_i} = 2c_i+\epsilon_i\ \ (1 \le i \le k)$$

οι αντίστοιχες ιδιοτιμές της διγραμμικής μορφής $\widetilde{q}$ και του ενδομορφισμού $\widetilde{c}$. Η διαφορά των κινδύνων γράφεται, λοιπόν

$$R_\phi(\theta,\sigma) - R_{\phi_\epsilon}(\theta,\sigma) = \frac{2}{\sigma^2}\mathbb{E}_{\theta,\sigma}\Big[h(t,u)\sum_{i=1}^{k}\widetilde{q_i}\widetilde{c_i}y_i(y_i-\theta_i)\Big] - \frac{1}{\sigma^2}\mathbb{E}_{\theta,\sigma}\Big[h^2(t,u)\sum_{i=1}^{k}\widetilde{q_i}\widetilde{c_i}^2 y_i^2\Big]$$

και εμφανίζεται σαν τη διαφορά των κινδύνων του ε.ε.τ ϕ_0 και του εκτιμητή με συρρικνωτή $\phi(y) = (id_\Theta - h(t,u)\widetilde{c})\phi_0(y)$ για τη συνάρτηση απωλείας που συνδέεται με την διγραμμική μορφή $\widetilde{q}$. Συνεπώς, σύμφωνα με το προηγούμενο Λήμμα, έχουμε το ζητούμενο αν:

$$\begin{aligned}
\sup_{t,u} tuh(t,u) &\le 2\ \frac{tr(v_\Theta\widetilde{q}\,\widetilde{c}) - 2\lambda_1 pgvp(v_\Theta\widetilde{q}\,\widetilde{c})}{pgvp({}^t\widetilde{c}\,\widetilde{q}\,\widetilde{c}\,b^-)}\ \frac{n-k}{n-k+2\lambda_2}\\
&= 2\ \frac{\sum_{i=1}^{k}\widetilde{q_i}\widetilde{c_i} - 2\lambda_1\max(\widetilde{q_i}\widetilde{c_i})}{\max\{\widetilde{c_i}q_i\widetilde{c_i}b_i^{-1}\}}\ \frac{n-k}{n-k+2\lambda_2}\\
&= 2\ \frac{\sum_{i=1}^{k} q_i\epsilon_i - 2\lambda_1\max(q_i\epsilon_i)}{\max\{q_i\epsilon_i(2c_i+\epsilon_i)b_i^{-1}\}}\ \frac{n-k}{n-k+2\lambda_2}
\end{aligned}$$

δηλαδή, αν ικανοποιείται η συνθήκη της εν λόγω πρότασης.

□

Επομένως για να μπορέσει ένας διαταραγμένος εκτιμητής να βελτιώσει τον ϕ πρέπει η $tuh(t,u)$ να είναι φραγμένη. Σε μια τέτοια περίπτωση, θα μπορούσαμε ξεκινώντας από c_i μικρά, να βελτιώνουμε τον εκτιμητή αυξάνοντας σταδιακά τις ιδιοτιμές κατά τρόπο ώστε το ανωτέρω φράξιμο να συνεχίζει να επαληθεύεται.

Παρατηρούμε ότι, για $c = 0$, επιστρέφουμε στο αποτέλεσμα του Λήμματος και ότι η τροποποίηση μιας ή δύο συνιστωσών δεν μας επιτρέπει να εφαρμόσουμε την Πρόταση 4.2.2 (επειδή $\lambda_1 \ge 1$ και σ'αυτή την περίπτωση το φράγμα βγαίνει αρνητικό).

4.2.2 Συρρικνωτές Κλίμακας

4.2.2(A).

Η εισαγωγή του ενδομορφισμού συρρίκνωσης μέσα στους εκτιμητές με συρρικνωτή από τους *Berger* και *Bock* βασίζεται στην περίπτωση όπου, μέσα στη θεωρούμενη αυτών βάση, ο πίνακας διασπορών-συνδιασπορών δεν είναι ο μοναδιαίος (σημειώνουμε επίσης οτι η επιλογή $c = (v_\Theta q)^{-1}$, εισάγεται φυσιολογικά στην περίπτωση που η συνάρτηση απωλείας είναι συνδεδεμένη με μία τετραγωνική μορφή άλλη από την $\overline{v_\Theta^{-1}}$).

Επειδή με την εισαγωγή του ενδομορφισμού συρρικνώσεως c, η συρρίκνωση είναι διαφορετική ακολουθώντας κάθε συντεταγμένη του $\phi_0(y)$ και ο προκύπτων εκτιμητής δεν τοποθετείται πλέον (όπως στην περίπτωση του εκτιμητή κλίμακας) επί της ευθείας που συνδέει τον εκλεγόμενο πόλο με τον ε.ε.τ, δεδομένου ενός εκτιμητή με συρρικνωτή *matriciel* θα εξετάσουμε υπό ποιές συνθήκες είναι προτιμότερο να χρησιμοποιούμε έναν συνδεδεμένο με αυτόν εκτιμητή κλίμακας(*scalaire*).Υποθέτουμε ότι ικανοποιούνται οι υποθέσεις ελέγχου της §4.2.1. Προκύπτει, τότε από την Πρόταση 4.2.2 (θέτοντας $\epsilon_i{=}\alpha - c_i$, $1{\le}i{\le}k$) ότι:

Πόρισμα 4.2.1

Ο εκτιμητής $\phi_s(y) = \Big(1 - \alpha h(t,u)\Big)\phi_0(y)$, *όπου* $\alpha \geq pgvp(c)$, *κυριαρχεί ομοιομόρφως τον* ϕ, *αν:*

$$\sup_{t,u} tuh(t,u) \leq 2\ \frac{\sum_{i=1}^{k} q_i(\alpha - c_i) - 2\lambda_1 \max q_i(\alpha - c_i)}{\max\Big(q_i b_i^{-1}(\alpha^2 - c_i^2)\Big)}\ \frac{n-k}{n-k+2\lambda_2}$$

Η συνθήκη $\alpha \geq pgvp(c)$, μας εξασφαλίζει ότι, για κάθε $i(1{\le}i{\le}k)$, το $\epsilon_i{\ge}0$.

4.2.2(Β). Επιλογή του α

Θεωρώντας δεδομένη την συνάρτηση h, θα υπάρχει λοιπόν ένας εκτιμητής με συρρικνωτή κλίμακας κινδύνου μικρότερου ή ίσου από εκείνο του ϕ, αν

$$\sup_{t,u} tuh(t,u) \leq \sup_{\alpha \geq pgvp(c)} \left[2\ \frac{\sum_{i=1}^{k} q_i(\alpha - c_i) - 2\lambda_1 \max q_i(\alpha - c_i)}{\max\Big(q_i b_i^{-1}(\alpha^2 - c_i^2)\Big)}\ \frac{n-k}{n-k+2\lambda_2} \right] \tag{4.3}$$

Επειδή γενικά, δεν είναι δυνατόν να μελετήσουμε την συνάρτηση του α τοιουτοτρόπως ορισμένη, θα θεωρήσουμε τη μερική περίπτωση όπου $q = b = v_\Theta^{-1}$. Έστω τότε

$$f(\alpha) = \Big[(k-2\lambda_1)\alpha - \Big(\sum_{i=1}^{k} c_i - 2\lambda_1 \min(c_i)\Big)\Big]\ \Big[\alpha^2 - (\min(c_i))^2\Big]^{-1} = \frac{(k-2\lambda_1)\alpha - \gamma}{\alpha^2 - m^2}$$

όπου $\gamma = \sum_{i=1}^{k} c_i {-} 2\lambda_1 m$ και $m = \min(c_i)$

Η διακρίνουσα της δευτεροβάθμιας εξίσωσης $\alpha^2 {-} \frac{2\gamma}{k-2\lambda_1}\alpha {+} m^2 = 0$(που ισοδυναμεί με την $f'(\alpha) = 0$) είναι: $\Delta = 4\frac{\gamma^2}{(k-2\lambda_1)^2} {-} 4m^2$, όπου $\gamma = \sum_{i=1}^{k} c_i {-} 2\lambda_1 m \geq (k{-}2\lambda_1)m$. Κατά συνέπεια, έχουμε ότι $\Delta{\ge}0$ και επομένως υπάρχουν δύο ρίζες α_1 και α_2 και τέτοιες ώστε $f'(\alpha_1){=}f'(\alpha_2){=}0$, με:

$$\alpha_1 = \frac{\gamma}{k-2\lambda_1} - \sqrt{\frac{\gamma^2}{(k-2\lambda_1)^2} - m^2} \geq 0 \quad \text{και} \quad \alpha_2 = \frac{\gamma}{k-2\lambda_1} + \sqrt{\frac{\gamma^2}{(k-2\lambda_1)^2} - m^2} \geq 0$$

Επίσης, παρατηρούμε ότι $\alpha_1 \leq m \leq \alpha_2$. Έχουμε, λοιπόν τον παρακάτω πίνακα μεταβολής της συναρτήσεως f:

α	0		α_1		m		α_2		$+\infty$
f	$\frac{\gamma}{m^2}$	$\searrow$	$f(\alpha_1)$	$\nearrow$	$+\infty \;\Vert\; -\infty$	$\nearrow$	$f(\alpha_2)$	$\searrow$	0

Θέτοντας, λοιπόν, $\alpha_0 = \sup\Big(\alpha_2, \max(c_i)\Big)$, υπάρχει εκτιμητής με συρρικνωτή κλίμακας κυριαρχώντας τον ϕ αν,

$$2f(\alpha_0)\frac{n-k}{n-k+2\lambda_2} \geq \sup_{t,u} tuh(t,u)$$

4.2.2(Γ).

Θεωρώντας μία συνάρτηση συρρικνώσεως h και συνδέοντας με αυτή έναν ενδομορφισμό c, θα πρέπει λοιπόν να επαληθεύσουμε ότι η συνθήκη (4.3) δεν ικανοποιείται. Όμως αυτή η συνθήκη δεν αρκεί για να βεβαιωθούμε ότι ο εκτιμητής που κατασκευάσαμε μ'αυτόν τον τρόπο κυριαρχεί όλους τους εκτιμητές με συρρικνωτή κλίμακας. Ωστόσο οι εκτιμητές με συρρικνωτή "*matriciel*" παρουσιάζουν ιδιαίτερο ενδιαφέρον όπως, για παράδειγμα, στην περίπτωση των εκτιμητών με τυχαίο πόλο (*Efron* και *Morris*[1975], *Manjogue* και *Rao*[1981]) ή στην περίπτωση της πολυσυγγραμικότητας ("*multicolinéarité*"), όταν οι εκτιμητές κλίμακας τείνουν όλοι προς τον ε.ε.τ.

4.2.3 Μία πλήρης υπο-κλάση

Η Πρόταση 4.2.2 μας επιτρέπει εξίσου να αποκλείσουμε ένα μέρος συναρτήσεων συρρίκνωσης αποδεικνύοντας οτι οι εκτιμητές που συνδέονται με αυτές κυριαρχούνται ομοιομόρφως. Θα δώσουμε παρακάτω, ακολουθώντας τον *C.Robert*[1987] ([41]), μία γενίκευση του αποτελέσματος των *Judge* και *Bock* (σελ.183-184) για τον εκτιμητή των *James-Stein* και του αποτελέσματος του Κεφ.1(Πρότ.1.4.5) για τους εκτιμητές της μορφής:

$$\phi(y) = \Big(id_\Theta - \rho\,\frac{1}{tu}\,c\Big)\phi_0(y)$$

Έχουμε λοιπόν:

Πόρισμα 4.2.2

Αν η h ικανοποιεί

$$\forall t,u \quad , \quad tuh(t,u) < \frac{\sum_{i=1}^{k} q_i c_i - 2\lambda_1 \max(q_i c_i)}{\max(q_i c_i^2 b_i^{-1})}\;\frac{n-k}{n-k+2\lambda_2}$$

τότε ο εκτιμητής που συνδέεται με την h και τον c είναι ομοιόμορφα κυριαρχημένος.

Απόδειξη

Θα κατασκευάσουμε έναν διαταραγμένο εκτιμητή ο οποίος βελτιώνει τον εκτιμητή:

$$\phi(y) = \Big(id_\Theta - h(t,u)c\Big)\phi_0(y)$$

Έστω $\alpha{>}1$ και $\epsilon_i{=}(\alpha-1)c_i$ $(1{\leq}i{\leq}k)$. Κατά συνέπεια, $\epsilon_i(2c_i+\epsilon_i){=}(\alpha^2-1)c_i^2$ και το φράγμα της Πρότασης 4.2.2 γίνεται:

$$2\,\frac{(\alpha-1)\sum_{i=1}^{k}q_ic_i-2\lambda_1(\alpha-1)\max(q_ic_i)}{(\alpha^2-1)\max(q_ic_i^2b_i^{-1})}\;\frac{n-k}{n-k+2\lambda_2}$$

ή ισοδύναμα

$$2\,\frac{1}{\alpha+1}\;\frac{\sum_{i=1}^{k}q_ic_i-2\lambda_1\max(q_ic_i)}{\max(q_ic_i^2b_i^{-1})}\;\frac{n-k}{n-k+2\lambda_2}$$

Η απεικόνιση $\alpha\rightsquigarrow(\alpha+1)^{-1}$ είναι μία φθίνουσα συνάρτηση του α με τιμή 1/2 στο $\alpha{=}1$. Αν η h ικανοποιεί τον περιορισμό του πορίσματος, λαμβάνοντας το α αρκετά κοντά στο 1, η συνθήκη της Προτ.4.2.2 θα ικανοποιείται, απ' όπου το αποτέλεσμα.

□

Παρατηρούμε λοιπόν ότι, ο εκτιμητής, βελτιώνοντας τον $\phi(y)=\Big(id_\Theta-h(t,u)c\Big)\phi_0(y)$, είναι ο $\widetilde{\phi}(y)=\Big(id_\Theta-\alpha h(t,u)c\Big)\phi_0(y)$.

4.3 ΕΝΔΟΜΟΡΦΙΣΜΟΙ ΜΕ ΑΡΝΗΤΙΚΕΣ ΙΔΙΟΤΙΜΕΣ

Στους *Cellier*, *Fourdrinier* και *Robert* καθώς και στην πλειονότητα των άρθρων μία βασική υπόθεση είναι ότι οι ιδιοτιμές του ενδομορφισμού c είναι μη-αρνητικές. Αυτός ο περιορισμός δικαιολογείται φυσικά από τον όρο αυτό καθ'αυτό της συρρίκνωσης: όταν μια ιδιοτιμή είναι αρνητική, ο εκτιμητής απομακρύνεται από τον πόλο στην αντίστοιχη διεύθυνση. Όμως, στο Κεφ.1, είδαμε επαρκείς συνθήκες για την κυριαρχία επί του ε.ε.τ οι οποίες εξακολουθούν να διατηρούνται αναλλοίωτες ακόμα και όταν κάποιες ιδιοτιμές του c είναι αρνητικές. Ιδιαιτέρως, εκεί είδαμε μία σαφή επαρκή συνθήκη όταν όλες οι ιδιοτιμές του c είναι αρνητικές(βλ.Παρατήρηση 1.4.4, Κεφ.1)

Όμως, στην περίπτωση που κάποιες ιδιοτιμές είναι αρνητικές δεν μπορούμε να καθορίσουμε τις συναρτήσεις συρρίκνωσης ώστε να επιτευχθεί μία ομοιόμορφη κυριαρχία επί του ε.ε.τ. Είναι, λοιπόν, προτιμότερο να περιοριστούμε στους εκτιμητές με ενδομορφισμό συρρίκνωσης που επιδέχεται μη αρνητικές ιδιοτιμές. Θα δούμε υπό ποιές προϋποθέσεις υπάρχει ένας τέτοιος εκτιμητής ο οποίος να βελτιώνει έναν εκτιμητή που δεν συρρικνώνεται προς όλες τις κατευθύνσεις.

Παρατηρούμε ότι η Πρόταση 4.2.2 ισχύει εξίσου για ενδομορφισμούς με αρνητικές ιδιοτιμές με την προϋπόθεση το ϵ να ικανοποιεί τη συνθήκη: $\epsilon_i+2c_i{>}\,0$ $(1{\leq}i{\leq}k)$, (τότε τα $\tilde{q}$ και $\tilde{c}$ της απόδειξης έχουν μη-αρνητικές ιδιοτιμές).

Υποθέτουμε στα παρακάτω ότι ισχύουν οι υποθέσεις του Κεφ.1 και επιπλέον ότι η h ικανοποιεί τις συνθήκες "ελέγχου" που ορίστηκαν στην §4.2.1 ((a) και (b)). Προκύπτει τότε:

Πόρισμα 4.3.3

Ο εκτιμητής ϕ του τύπου (4.2) κυριαρχείται από τον εκτιμητή

$$\phi_d(y)=\Big(id_\Theta-h(t,u)d\Big)\phi_0(y)\ \ \text{όπου}\ \ d_i\geq|c_i|\ \ (1\leq i\leq k)$$

αν

$$\sup_{t,u}tuh(t,u)\leq 2\,\frac{\sum_{i=1}^{k}q_i(d_i-c_i)-2\lambda_1\max q_i(d_i-c_i)}{\max q_ib_i^{-1}(d_i^2-c_i^2)}\;\frac{n-k}{n-k+2\lambda_2}$$

Το παραπάνω πόρισμα είναι άμεση συνέπεια της Πρότασης 4.2.2 λαμβάνοντας $\epsilon_i = d_i - c_i \geq 0$ $(1 \leq i \leq k)$ οπότε ο $c_\epsilon = d$.

Λαμβάνουμε εξίσου το παρακάτω αποτέλεσμα, που επιβάλλει λιγότερους περιορισμούς επί της συνάρτησης συρρίκνωσης(διατηρούμε μόνο την (a)).

Πρόταση 4.3.3

Ο εκτιμητής ϕ, που συνδέεται με την h και τον c, κυριαρχείται από τον εκτιμητή

$$\phi^+(y) = \Big(id_\Theta - h(t,u)d\Big)\phi_0(y)\ , \quad \text{όπου } \ d_i = |c_i| \ \ (1 \leq i \leq k)$$

αν $\displaystyle\sum_{i \in J} q_i|c_i| - 2\lambda_1 \max_{i \in J} q_i|c_i| \geq 0\ ,$ *όπου* $J = \Big\{i \in \{1, ..., k\} : c_i < 0\Big\}$.

Απόδειξη

Έχουμε:

$$\begin{aligned}
&R_\phi(\theta,\sigma) - R_{\phi^+}(\theta,\sigma)\\
&= \frac{1}{\sigma^2}\mathbb{E}_{\theta,\sigma}\Big[h^2(t,u)\Big(\overline{q}(c(\phi_0(y))) - \overline{q}(d(\phi_0(y)))\Big)\Big] + \frac{2}{\sigma^2}\mathbb{E}_{\theta,\sigma}\Big[h(t,u)q\Big(\phi_0(y) - \theta, d(\phi_0(y)) - c(\phi_0(y))\Big)\Big]\\
&= \frac{1}{\sigma^2}\mathbb{E}_{\theta,\sigma}\Big[h^2(t,u)\sum_{i=1}^{k}\underbrace{(c_i^2 - d_i^2)}_{=\,0}y_i^2 q_i\Big] + \frac{2}{\sigma^2}\mathbb{E}_{\theta,\sigma}\Big[h(t,u)\sum_{i=1}^{k}(d_i - c_i)q_i y_i(y_i - \theta_i)\Big]\\
&= \frac{4}{\sigma^2}\mathbb{E}_{\theta,\sigma}\Big[h(t,u)\sum_{i \in J} q_i|c_i|y_i(y_i - \theta_i)\Big]
\end{aligned}$$

επειδή, δοθέντος ότι $d_i = |c_i|$ για κάθε $i \in \{1, ..., k\}$, $d_i^2 = c_i^2$ και

$$\sum_{i=1}^{k} q_i(d_i - c_i)y_i(y_i - \theta_i) = \sum_{i: c_i < 0} q_i(d_i - c_i)y_i(y_i - \theta_i) = 2\sum_{i \in J} q_i|c_i|y_i(y_i - \theta_i)$$

Ολοκληρώνοντας κατά μέρη(χρησιμοποιώντας το Λήμμα του $Stein$ όπως στην απόδειξη της Πρότ.1.3.2, Κεφ.1 για τον όρο $A_\phi(\theta,\sigma)$), λαμβάνουμε ότι:

$$R_\phi(\theta,\sigma) - R_{\phi^+}(\theta,\sigma) = 4\ \mathbb{E}_{\theta,\sigma}\Big[h(t,u)\sum_{i \in J} q_i|c_i| + 2h_1'(t,u)\sum_{i \in J} q_i b_i|c_i|y_i^2\Big]$$

Όμως επειδή $t^{\lambda_1}h(t,u)$ είναι αύξουσα συνάρτηση του t, για κάθε u, (λόγω του (a)) έχουμε ότι $\lambda_1 t^{\lambda_1 - 1}h(t,u) + t^{\lambda_1}h_1'(t,u) \geq 0$, δηλαδή $\lambda_1 h(t,u) + t h_1'(t,u) \geq 0$, απ' όπου προκύπτει ότι:

$$h_1'(t,u) \geq -\lambda_1 \frac{1}{t}\ h(t,u)$$

Συνεπώς,

$$\begin{aligned}
R_\phi(\theta,\sigma) - R_{\phi^+}(\theta,\sigma) \ &\geq\ 4\ \mathbb{E}_{\theta,\sigma}\Big[h(t,u)\Big(\sum_{i \in J} q_i|c_i| - 2\lambda_1\frac{1}{t}\sum_{i \in J} q_i b_i|c_i|y_i^2\Big)\Big]\\
&\geq\ 4\ \mathbb{E}_{\theta,\sigma}\Big[h(t,u)\Big(\sum_{i \in J} q_i|c_i| - 2\lambda_1 \max_{i \in J}(q_i|c_i|)\underbrace{\frac{1}{t}\sum_{i \in J} b_i y_i^2}_{\leq\, 1}\Big)\Big]\\
&\geq\ 4\ \mathbb{E}_{\theta,\sigma}\Big[h(t,u)\Big(\sum_{i \in J} q_i|c_i| - 2\lambda_1 \max_{i \in J}(q_i|c_i|)\Big)\Big]
\end{aligned}$$

το οποίο είναι μεγαλύτερο ή ίσο του μηδενός αν ισχύει η συνθήκη της Πρότασης.

□

Μία άμεση συνέπεια της Πρότασης 4.3.3 είναι:

Πόρισμα 4.3.4

Μία ικανή συνθήκη ώστε ο εκτιμητής $\phi(y) = \Big(1 + \alpha h(t,u)\Big)\phi_0(y)$, *όπου* $\alpha \in \mathbb{R}_+$, *να κυριαρχείται από τον* $\phi^+(y) = \Big(1 - \alpha h(t,u)\Big)\phi_0(y)$ *είναι:* $k - 2\lambda_1 \geq 0$.

Πράγματι, λαμβάνοντας $c = -\alpha id_\Theta$ και $d = \alpha id_\Theta$, ικανοποιούνται προφανώς οι συνθήκες της Πρότασης 4.3.3:

$$\forall i \;\; d_i = |c_i| \; , \;\; \text{αφού} \;\; d_i = \alpha \;\; \text{και} \;\; c_i = -\alpha$$

και
$$\sum_{i\in J} q_i|c_i| - 2\lambda_1 \max_{i\in J}(q_i|c_i|) = \alpha\Big(\sum_{i=1}^{k} q_i - 2\lambda_1 \max_{1\leq i\leq k} q_i\Big) \leq \alpha(k-2\lambda_1) \max_{1\leq i\leq k} q_i$$

Συνεπώς απαραίτητη προϋπόθεση για να ισχύει η συνθήκη της Πρότ.4.3.3 είναι το $k-2\lambda_1 \geq 0$.

4.4 ΜΙΑ ΙΔΙΟΤΗΤΑ ΤΟΥ ΕΚΤΙΜΗΤΗ ΤΩΝ JAMES – STEIN

Έχουμε αποδείξει στην §4.2.2 ότι, για μία συγκεκριμένη κλάση εκτιμητών με συρρικνωτή, υπάρχουν εκτιμητές με συρρικνωτή κλίμακας που τους κυριαρχούν (με την έννοια του κινδύνου). Θα αποδείξουμε την κυριαρχία του εκτιμητή των $James-Stein$ για ένα κριτήριο πιο χονδροειδές από αυτό του τετραγωνικού κινδύνου: την κυριαρχία με την έννοια του *εκτιμητή ενός άνω φράγματος του κινδύνου* ή *"$m-$κυριαρχία"*.

Ακολουθώντας τους $Efron$ και $Morris$[1976], πολυάριθμοι συγγραφείς χρησιμοποιούν έναν αμερόληπτο εκτιμητή του κινδύνου για να εξασφαλίσουν ικανές συνθήκες για την κυριαρχία επί του ε.ε.τ. Από αυτόν τον αμερόληπτο εκτιμητή, ο οποίος είναι μοναδικός για τον κανονικό νόμο, οι $Moore$ και $Brooks$[1978] συμπεραίνουν ένα κριτήριο σύγκρισης των εκτιμητών ονομαζόμενο *κριτήριο της σύγκρισης με την έννοια του εκτιμητή του κινδύνου*. Συγκεκριμένα, αν $R_\phi(\theta,\sigma) = \mathbb{E}_{\theta,\sigma}[\Psi(\phi;y)]$, ο ϕ θα κυριαρχεί τον ϕ' με την έννοια του κριτηρίου αυτού αν $\Psi(\phi;y) \leq \Psi(\phi';y)$, $\forall y$. Αυτοί αποδεικνύουν ότι ο εκτιμητής των $James - Stein$ δεν κυριαρχείται με την έννοια του κριτηρίου αυτού, από κανέναν εκτιμητή με συρρικνωτή κλίμακας. Θα μπορούσε κανείς να γενικεύσει το αποτέλεσμα τους σε όλους τους εκτιμητές με συρρικνωτή h τέτοιο ώστε η $tuh(t,u)$ να είναι φραγμένη.

Θα συμπεράνουμε από αυτό το κριτήριο έναν "χονδροειδή" τρόπο σύγκρισης, θεωρώντας το άνω φράγμα του Ψ το οποίο θα χρησιμοποιηθεί για να προκύψουν ικανές συνθήκες για την κυριαρχία επί του ε.ε.τ. Θεωρούμε ότι ισχύουν οι υποθέσεις του Κεφ.1 και θέτουμε $q = b = v_\Theta^{-1}$. Από την Πρόταση 1.3.2 του Κεφ.1, έχουμε ότι:

$$\begin{aligned} R_{\phi_0}(\theta,\sigma) - R_\phi(\theta,\sigma) \;\; &= \;\; 2\,\mathbb{E}_{\theta,\sigma}\Big[\sum_{i=1}^{k} c_i h(t,u) + 2h_1'(t,u)\sum_{i=1}^{k} c_i y_i^2\Big] \\ &\quad - \mathbb{E}_{\theta,\sigma}\Big[\sum_{i=1}^{k} c_i^2 y_i^2 \Big\{-4h(t,u)h_2'(t,u)\frac{u^2}{n-k} + \frac{n-k-2}{n-k}uh^2(t,u)\Big\}\Big] \end{aligned}$$

Αν υποθέσουμε ότι οι συνθήκες του "ελέγχου" ικανοποιούνται, οπότε προκύπτουν αντίστοιχα οι σχέσεις:

$$h_1'(t,u) \geq -\lambda_1 \frac{1}{t} h(t,u) \quad \text{και} \quad h_2'(t,u) \geq -\lambda_2 \frac{1}{u} h(t,u) \;\; ,$$

λαμβάνουμε ότι:

$$R_{\phi_0}(\theta,\sigma) - R_{\phi}(\theta,\sigma) \;\geq\; 2\,\mathbb{E}_{\theta,\sigma}\Big[h(t,u)\Big\{\sum_{i=1}^{k} c_i - 2\lambda_1 \max(c_i)\Big\}\Big] - \mathbb{E}_{\theta,\sigma}\Big[tuh^2(t,u)\frac{n-k-2+4\lambda_2}{n-k}\max(c_i^2)\Big]$$

Και αν $\beta = \sup_{t,u} tuh(t,u)$, συμπεραίνουμε ότι :

$$R_{\phi_0}(\theta,\sigma) - R_{\phi}(\theta,\sigma) \geq \mathbb{E}_{\theta,\sigma}\Big[h(t,u)\Big\{2\Big(\sum_{i=1}^{k} c_i - 2\lambda_1 \max(c_i)\Big) - \beta\frac{n-k-2+4\lambda_2}{n-k}\max(c_i^2)\Big\}\Big]$$

Ονομάζοντας *εκτιμητή του άνω φράγματος του κινδύνου* συνδεδεμένο με τα λ_1, λ_2, τη συνάρτηση:

$$\xi_{\lambda_1,\lambda_2}(\phi;y) = h(t,u)\Big\{2\Big(\sum_{i=1}^{k} c_i - 2\lambda_1 \max(c_i)\Big) - \beta\frac{n-k-2+4\lambda_2}{n-k}\max(c_i^2)\Big\}$$

θα λέμε ότι ο ϕ *$m-$κυριαρχεί* τον ϕ' αν

$$\xi_{\lambda_1,\lambda_2}(\phi;y) \geq \xi_{\lambda_1,\lambda_2}(\phi';y) \quad , \quad \forall y$$

Αποδεικνύουμε τότε:

Πρόταση 4.4.4

Ο εκτιμητής των $James-Stein$, $m-$κυριαρχεί τους εκτιμητές με συρρικνωτή που συνδέονται με μία συνάρτηση συρρικνώσεως h τέτοια ώστε η $h(\,\cdot\,,u)$ να είναι ελεγχόμενη σε ύψος λ_1 για κάθε u, η $h(t,\cdot)$ να είναι ελεγχόμενη σε ύψος λ_2 και η $tuh(t,u)$ να είναι ομοιόμορφα φραγμένη.

Απόδειξη

Θεωρώντας τον $\xi_{\lambda_1,\lambda_2}(\phi;y)$ παρατηρούμε ότι:

$$2\Big(\sum_{i=1}^{k} c_i - 2\lambda_1 \max(c_i)\Big) - \beta\frac{n-k-2+4\lambda_2}{n-k}\max(c_i^2)$$

μεγιστοποιείται για $c_1 = \ldots = c_k = \dfrac{(k-2\lambda_1)(n-k)}{\beta(n-k-2+4\lambda_2)}$ και για $\lambda_1 = \lambda_2 = 1$ ενώ $h(t,u) \leq \dfrac{\beta}{tu}$. Επομένως

$$\begin{aligned}
\xi_{\lambda_1,\lambda_2}(\phi;y) &\leq \frac{\beta}{tu}\Big\{2k\frac{(k-2)(n-k)}{\beta(n-k+2)} - 4\frac{(k-2)(n-k)}{\beta(n-k+2)} - \beta\frac{n-k+2}{n-k}\;\frac{(k-2)^2(n-k)}{\beta^2(n-k+2)^2}\Big\} \\
&= \frac{\beta}{tu}\Big\{2\frac{(k-2)^2(n-k)}{\beta(n-k+2)} - \frac{(k-2)^2(n-k)}{\beta(n-k+2)}\Big\} \\
&= \frac{1}{tu}\;\frac{(k-2)^2(n-k)}{n-k+2} \\
&= \xi_{1,1}(\phi_{JS};y)
\end{aligned}$$

όπου $\xi_{1,1}(\phi_{JS}; y)$ είναι ο εκτιμητής του άνω φράγματος του κινδύνου του εκτιμητή των $James-Stein$:

$$\phi_{JS}(y) = \Big(1 - \frac{(n-k)(k-2)}{n-k+2}\ \frac{1}{tu}\Big)\phi_0(y)$$

για τον οποίον $h(t,u) = \dfrac{(n-k)(k-2)}{n-k+2}\ \dfrac{1}{tu}$ και $c = id_\Theta$ (επίσης στην περίπτωση αυτή $\beta = \dfrac{(n-k)(k-2)}{n-k+2}$).

□

4.5 ΤΡΟΠΟΠΟΙΗΣΗ ΤΗΣ ΣΥΝΑΡΤΗΣΗΣ ΣΥΡΡΙΚΝΩΣΗΣ

Έως τώρα στο κεφάλαιο αυτό θεωρήσαμε τροποποιήσεις του ενδομορφισμού συρρίκνωσης υπό μία δεδομένη συνάρτηση συρρίκνωσης h. Σ'αυτή την παράγραφο, θα δώσουμε ένα κριτήριο για την κυριαρχία επί ενός εκτιμητή με συρρικνωτή, τροποποιώντας την συνάρτηση συρρικνώσεως. Για να απλοποιήσουμε τους υπολογισμούς, υποθέτουμε ότι η b είναι θετικά ορισμένη και οι συναρτήσεις συρρικνώσεως είναι διαφορίσιμες.

Προκύπτει τότε, υπό τις υποθέσεις (H_1), (H_2) και (H_4) του Κεφ.1, ότι:

Πρόταση 4.5.5

Ο εκτιμητής $\phi_h(y)=\Big(id_\Theta - h(t,u)c\Big)\phi_0(y)$ *κυριαρχεί τον εκτιμητή* $\phi_l(y)=\Big(id_\Theta - l(t,u)c\Big)\phi_0(y)$ *ομοιομόρφως στο* (θ,σ), *αν*

(i) υπάρχει λ_1 *τέτοιο ώστε η* $t \rightsquigarrow t^{\lambda_1}(h(t,u) - l(t,u))$ *να είναι αύξουσα συνάρτηση του* t, *για κάθε* u

(ii) υπάρχει λ_2 *τέτοιο ώστε η* $u \rightsquigarrow u^{\lambda_2}(h(t,u) - l(t,u))$ *να είναι αύξουσα συνάρτηση του* u, *για κάθε* t

(iii) αν $h(t,u) > l(t,u)$ *τότε*

$$tu(h+l)(t,u) \leq 2\ \frac{tr(v_\Theta qc) - 2\lambda_1 pgvp(v_\Theta qc)}{pgvp({}^t cqcb^{-1})}\ \frac{n-k}{n-k-2+2\lambda_2}$$

(iv) αν $h(t,u) < l(t,u)$ *τότε*

$$tu(h+l)(t,u) \geq 2\ \frac{tr(v_\Theta qc) - 2\lambda_1 ppvp(v_\Theta qc)}{ppvp({}^t cqcb^{-1})}\ \frac{n-k}{n-k-2+2\lambda_2}$$

όπου $ppvp$ *συμβολίζει τη μικρότερη ιδιοτιμή του αντίστοιχου ενδομορφισμού.*

Απόδειξη

Χρησιμοποιώντας την Πρότ.1.3.2(Κεφ.1), η διαφορά των κινδύνων μεταξύ των δύο εκτιμητών γράφεται:

$$\begin{aligned}
&R_{\phi_l}(\theta,\sigma) - R_{\phi_h}(\theta,\sigma)\\
&= 2\mathbb{E}_{\theta,\sigma}\Big[tr(v_\Theta qc)\Big(h(t,u) - l(t,u)\Big) + 2\overline{bv_\Theta qc}(\phi_0(y))\Big(h_1'(t,u) - l_1'(t,u)\Big)\Big]\\
&\quad -\mathbb{E}_{\theta,\sigma}\Big[\overline{q}\Big(c(\phi_0(y))\Big)\Big\{-4\frac{u^2}{n-k}\Big(h(t,u)h_2'(t,u) - l(t,u)l_2'(t,u)\Big) + \frac{n-k-2}{n-k}u\Big(h^2(t,u) - l^2(t,u)\Big)\Big\}\Big]
\end{aligned}$$

$$= 2\mathbb{E}_{\theta,\sigma}\Big[tr(v_\Theta qc)\Big(h(t,u)-l(t,u)\Big) + 2\sum_{i=1}^{k} q_i c_i b_i y_i^2 \Big(h_1'(t,u)-l_1'(t,u)\Big)\Big]$$

$$-\mathbb{E}_{\theta,\sigma}\Big[\sum_{i=1}^{k} q_i c_i^2 y_i^2 \Big\{-4\frac{u^2}{n-k}\Big(h(t,u)h_2'(t,u)-l(t,u)l_2'(t,u)\Big) + \frac{n-k-2}{n-k}u\Big(h^2(t,u)-l^2(t,u)\Big)\Big\}\Big]$$

Όμως, λόγω του (i), έχουμε ότι:

$$\lambda_1 t^{\lambda_1 - 1}\Big(h(t,u)-l(t,u)\Big) + t^{\lambda_1}\Big(h_1'(t,u)-l_1'(t,u)\Big) \geq 0$$

δηλαδή,

$$\lambda_1\Big(h(t,u)-l(t,u)\Big) + t\Big(h_1'(t,u)-l_1'(t,u)\Big) \geq 0$$

απ᾽ όπου λαμβάνουμε:

$$h_1'(t,u)-l_1'(t,u) \geq \lambda_1 \frac{1}{t}\Big(h(t,u)-l(t,u)\Big) \tag{4.4}$$

Ομοίως, λόγω του (ii), έχουμε:

$$\lambda_2 u^{\lambda_2 - 1}\Big(h^2(t,u)-l^2(t,u)\Big) + u^{\lambda_2}\Big(2h(t,u)h_2'(t,u)-2l(t,u)l_2'(t,u)\Big) \geq 0$$

δηλαδή,

$$\lambda_2\Big(h^2(t,u)-l^2(t,u)\Big) + 2u\Big(h(t,u)h_2'(t,u)-l(t,u)l_2'(t,u)\Big) \geq 0$$

απ᾽ όπου λαμβάνουμε:

$$h(t,u)h_2'(t,u)-l(t,u)l_2'(t,u) \geq \frac{\lambda_2}{2}\ \frac{1}{u}\Big(h^2(t,u)-l^2(t,u)\Big) \tag{4.5}$$

Συνεπώς, από τις σχέσεις (4.4) και (4.5), έχουμε ότι:

$$\begin{aligned}
& R_{\phi_l}(\theta,\sigma) - R_{\phi_h}(\theta,\sigma) \\
& \geq 2\ \mathbb{E}_{\theta,\sigma}\Big[\Big(h(t,u)-l(t,u)\Big)\Big(tr(v_\Theta qc) - 2\lambda_1\frac{1}{t}\sum_{i=1}^{k} q_i c_i b_i y_i^2\Big)\Big] \\
& \quad -\mathbb{E}_{\theta,\sigma}\Big[\Big(h^2(t,u)-l^2(t,u)\Big)u\sum_{i=1}^{k} q_i c_i^2 y_i^2\ \frac{n-k-2+2\lambda_2}{n-k}\Big] \\
& = \mathbb{E}_{\theta,\sigma}\Bigg[\Big(h(t,u)-l(t,u)\Big)\Big\{2\Big(tr(v_\Theta qc) - 2\lambda_1\frac{1}{t}\sum_{i=1}^{k} q_i c_i b_i y_i^2\Big) \\
& \qquad -\Big(h(t,u)+l(t,u)\Big)u\sum_{i=1}^{k} q_i c_i^2 y_i^2\ \frac{n-k-2+2\lambda_2}{n-k}\Big\}\Bigg]
\end{aligned}$$

Διακρίνουμε λοιπόν τις περιπτώσεις:

- Αν $h(t,u) > l(t,u)$ τότε, λόγω του (iii), έχουμε:

$$\Big(h(t,u)+l(t,u)\Big)u\sum_{i=1}^{k} q_i c_i^2 y_i^2\ \frac{n-k-2+2\lambda_2}{n-k}$$

$$
\begin{aligned}
&\leq 2\ \frac{1}{t}\ \frac{tr(v_\Theta qc) - 2\lambda_1 pgvp(v_\Theta qc)}{pgvp({}^t cqcb^{-1})} \sum_{i=1}^{k} q_i c_i^2 y_i^2 \\
&\leq 2\ \frac{1}{t}\ \frac{tr(v_\Theta qc) - 2\lambda_1 \max(q_i c_i)}{\max(q_i c_i^2 b_i^{-1})}\ \max(q_i c_i^2 b_i^{-1}) \sum_{i=1}^{k} b_i y_i^2 \\
&= 2\Big(tr(v_\Theta qc) - 2\lambda_1 \max(q_i c_i)\Big) \\
&= 2\Big(tr(v_\Theta qc) - 2\lambda_1 \frac{1}{t}\ \max(q_i c_i) \sum_{i=1}^{k} b_i y_i^2\Big) \\
&\leq 2\Big(tr(v_\Theta qc) - 2\lambda_1 \frac{1}{t} \sum_{i=1}^{k} q_i c_i b_i y_i^2\Big)
\end{aligned}
$$

από το οποίο συμπεραίνουμε ότι, $R_{\phi_l}(\theta,\sigma) - R_{\phi_h}(\theta,\sigma) \geq 0$.

- Αν $h(t,u) < l(t,u)$ τότε, λόγω του (iv), έχουμε:

$$
\begin{aligned}
&\Big(h(t,u) + l(t,u)\Big) u \sum_{i=1}^{k} q_i c_i^2 y_i^2\ \frac{n-k-2+2\lambda_2}{n-k} \\
&\geq 2\ \frac{1}{t}\ \frac{tr(v_\Theta qc) - 2\lambda_1 ppvp(v_\Theta qc)}{ppvp({}^t cqcb^{-1})} \sum_{i=1}^{k} q_i c_i^2 y_i^2 \\
&\geq 2\ \frac{1}{t}\ \frac{tr(v_\Theta qc) - 2\lambda_1 \min(q_i c_i)}{\min(q_i c_i^2 b_i^{-1})}\ \min(q_i c_i^2 b_i^{-1}) \sum_{i=1}^{k} b_i y_i^2 \\
&= 2\Big(tr(v_\Theta qc) - 2\lambda_1 \min(q_i c_i)\Big) \\
&= 2\Big(tr(v_\Theta qc) - 2\lambda_1 \frac{1}{t}\ \min(q_i c_i) \sum_{i=1}^{k} b_i y_i^2\Big) \\
&\geq 2\Big(tr(v_\Theta qc) - 2\lambda_1 \frac{1}{t} \sum_{i=1}^{k} q_i c_i b_i y_i^2\Big)
\end{aligned}
$$

απ᾽ όπου συμπεραίνουμε ότι και σ᾽ αυτήν την περίπτωση, $R_{\phi_l}(\theta,\sigma) - R_{\phi_h}(\theta,\sigma) \geq 0$.

$\square$

4.6 ΒΕΛΤΙΩΣΗ ΥΠΟ "ΚΟΛΟΒΩΣΗ" ΤΩΝ ΕΚΤΙΜΗΤΩΝ ΜΕ ΣΥΡΡΙΚΝΩΣΗ

Μία ενδιαφέρουσα "κλασσική" περίπτωση βελτίωσης εκτιμητών με συρρικνωτή, ενεργώντας επί της συνάρτησης συρρικνώσεως, αποτελούν οι "κολοβοί" (*tronqués*) εκτιμητές κλίμακας.

4.6.1 "Κόψιμο" εκ των κάτω

Περιοριζόμαστε στην περίπτωση όπου $q = b = v_\Theta^{-1}$ και $c = id_\Theta$, δηλαδή θεωρούμε τους εκτιμητές της μορφής:

$$
\phi(y) = \left(1 - h\Big(\overline{v_\Theta^{-1}}(\phi_0(y)), \frac{1}{s^2(y)}\Big)\right)\phi_0(y) \tag{4.6}
$$

οι οποίοι συγκρίνονται με τον ε.ε.τ υπό την συνήθη τετραγωνική απώλεια:

$$\sigma^{-2}\overline{v_{\Theta}^{-1}}(\widehat{\theta}-\theta) \tag{4.7}$$

Η βελτίωση την οποία επιφέρουν οι εκτιμητές αυτοί αναφορικά με τον ε.ε.τ ϕ_0 οφείλεται στο γεγονός ότι η συνάρτηση συρρικνώσεως h συρρικνώνει την εκτίμηση $\phi_0(y)$ προς τον πόλο μηδέν. Βεβαίως αυτή η συρρίκνωση πραγματοποιείται μόνο όταν

$$h\Big(\overline{v_{\Theta}^{-1}}(\phi_0(y)), \frac{1}{s^2(y)}\Big) \geq 0$$

αυτός είναι ο λόγος για τον οποίο σε ορισμένες εργασίες επιβάλλεται στην h να λαμβάνει τιμές θετικές ή μηδέν (περίπτωση την οποία θεωρήσαμε στα προηγούμενα κεφάλαια).

Φαίνεται λοιπόν λογικό, αν διαθέτουμε έναν εκτιμητή με συνάρτηση συρρικνώσεως h, η οποία μπορεί να πάρει και αρνητικές τιμές, να θεωρήσουμε τον εκτιμητή ο οποίος θα προκύψει από αυτόν αντικαθιστώντας την h από το θετικό της μέρος $h^+(=sup(0,h))$. Ορίζεται τότε, ένας νέος εκτιμητής του τύπου (4.6), ο οποίος βελτιώνει τον αρχικό εκτιμητή.

Ορισμός

Μία συνάρτηση $f:\mathbb{R}_+ \to \mathbb{R}_+$ θα λέμε ότι *ανήκει στην κλάση* $\mathcal{C}$ εάν αυτή ικανοποιεί τη συνθήκη:

$$\forall k \in \mathbb{N}_0 \ , \quad \int_{\mathbb{R}_+} f(w)(w-k)\mathcal{X}_k^2(dw) \geq 0$$

Παράδειγμα

Για κάθε $r \geq 0$, η συνάρτηση $w \rightsquigarrow w^r$, ανήκει στην κλάση $\mathcal{C}$.
Πράγματι,

$$\begin{aligned}
\int_{\mathbb{R}_+} w^r(w-k)\mathcal{X}_k^2(dw) &= \int_{\mathbb{R}_+} w^{r+1}\mathcal{X}_k^2(dw) - k\int_{\mathbb{R}_+} w^r\mathcal{X}_k^2(dw) \\
&= 2^{r+1}\frac{\Gamma(\frac{k}{2}+r+1)}{\Gamma(\frac{k}{2})} - k\,2^r\frac{\Gamma(\frac{k}{2}+r)}{\Gamma(\frac{k}{2})} \\
&= 2^{r+1}\Big(\frac{k}{2}+r\Big)\frac{\Gamma(\frac{k}{2}+r)}{\Gamma(\frac{k}{2})} - k\,2^r\frac{\Gamma(\frac{k}{2}+r)}{\Gamma(\frac{k}{2})} \\
&= 2^{r+1}r\frac{\Gamma(\frac{k}{2}+r)}{\Gamma(\frac{k}{2})}
\end{aligned}$$

Πρόταση 4.6.6

Υποθέτουμε ότι η συνάρτηση h, ορισμένη επί του $\mathbb{R}_+\times\mathbb{R}_+$, είναι τέτοια ώστε, για κάθε u, η συνάρτηση

$$w \rightsquigarrow h^-(\sigma^2 w, u)$$

όπου $h^- = sup(0,-h)$, να ανήκει στην κλάση $\mathcal{C}$. Τότε, για κάθε θ, ο εκτιμητής του τύπου (4.6) βελτιώνεται από τον εκτιμητή:

$$\phi_{h^+}(y) = \bigg(1 - h^+\Big(\overline{v_{\Theta}^{-1}}(\phi_0(y)), \frac{1}{s^2(y)}\Big)\bigg)\phi_0(y)$$

με την εξής έννοια:

$$\forall\theta,\sigma \quad R_\phi(\theta,\sigma)\geq R_{\phi_{h^+}}(\theta,\sigma)$$

Απόδειξη

Η διαφορά των κινδύνων των δύο εκτιμητών, γράφεται:

$$\begin{aligned}
&R_\phi(\theta,\sigma)-R_{\phi_{h^+}}(\theta,\sigma)\\
&=\frac{1}{\sigma^2}\int_E\Bigg\{\overline{v_\Theta^{-1}}\Big(\phi_0(y)-\theta-h\Big(\overline{v_\Theta^{-1}}(\phi_0(y)),\frac{1}{s^2(y)}\Big)\phi_0(y)\Big)\\
&\qquad-\overline{v_\Theta^{-1}}\Big(\phi_0(y)-\theta-h^+\Big(\overline{v_\Theta^{-1}}(\phi_0(y)),\frac{1}{s^2(y)}\Big)\phi_0(y)\Big)\Bigg\}N_E(\theta,\sigma^2 v;dy)\\
&=\frac{1}{\sigma^2}\int_{\Theta\times\Theta^\perp}\Bigg\{\overline{v_\Theta^{-1}}\Big(x-\theta-h\Big(\overline{v_\Theta^{-1}}(x),\frac{n-k}{\overline{v_{\Theta^\perp}^{-1}}(z)}\Big)x\Big)\\
&\qquad-\overline{v_\Theta^{-1}}\Big(x-\theta-h^+\Big(\overline{v_\Theta^{-1}}(x),\frac{n-k}{\overline{v_{\Theta^\perp}^{-1}}(z)}\Big)x\Big)\Bigg\}N_\Theta(\theta,\sigma^2 v_\Theta;dx)N_{\Theta^\perp}(0,\sigma^2 v_{\Theta^\perp};dz)\\
&=\frac{1}{\sigma^2}\int_{\mathbb{R}^k\times\mathbb{R}^{n-k}}\Bigg\{\Big\|x-\theta-h\Big(\|x\|^2,\frac{n-k}{\|z\|^2}\Big)x\Big\|^2\\
&\qquad-\Big\|x-\theta-h^+\Big(\|x\|^2,\frac{n-k}{\|z\|^2}\Big)x\Big\|^2\Bigg\}N_k(\theta,\sigma^2 I_k;dx)N_{n-k}(0,\sigma^2 I_{n-k};dz) \qquad (4.8)\\
&=\frac{1}{\sigma^2}(A+2B)
\end{aligned}$$

όπου

$$A=\int_{\mathbb{R}^k\times\mathbb{R}^{n-k}}\Big\{h^2\Big(\|x\|^2,\frac{n-k}{\|z\|^2}\Big)-h^{+2}\Big(\|x\|^2,\frac{n-k}{\|z\|^2}\Big)\Big\}\|x\|^2N_k(\theta,\sigma^2 I_k;dx)N_{n-k}(0,\sigma^2 I_{n-k};dz)$$

και

$$B=\int_{\mathbb{R}^k\times\mathbb{R}^{n-k}}h^-\Big(\|x\|^2,\frac{n-k}{\|z\|^2}\Big)\,{}^t(x-\theta)xN_k(\theta,\sigma^2 I_k;dx)N_{n-k}(0,\sigma^2 I_{n-k};dz)$$

Πράγματι, η υπό ολοκλήρωση ποσότητα στην (4.8),λαμβάνοντας υπόψιν ότι $h=h^+-h^-$, γράφεται:

$$\begin{aligned}
&\Big\|x-\theta-h\Big(\|x\|^2,\frac{n-k}{\|z\|^2}\Big)x\Big\|^2-\Big\|x-\theta-h^+\Big(\|x\|^2,\frac{n-k}{\|z\|^2}\Big)x\Big\|^2\\
&=\Big(h^2\Big(\|x\|^2,\frac{n-k}{\|z\|^2}\Big)-h^{+2}\Big(\|x\|^2,\frac{n-k}{\|z\|^2}\Big)\Big)\|x\|^2\\
&\quad+2\Big(h^+\Big(\|x\|^2,\frac{n-k}{\|z\|^2}\Big)-h\Big(\|x\|^2,\frac{n-k}{\|z\|^2}\Big)\Big)\,{}^t(x-\theta)x\\
&=\Big(h^2\Big(\|x\|^2,\frac{n-k}{\|z\|^2}\Big)-h^{+2}\Big(\|x\|^2,\frac{n-k}{\|z\|^2}\Big)\Big)\|x\|^2+2h^-\Big(\|x\|^2,\frac{n-k}{\|z\|^2}\Big)\,{}^t(x-\theta)x
\end{aligned}$$

το οποίο δικαιολογεί την τελευταία έκφραση της διαφοράς των κινδύνων.

Είναι προφανές ότι $A \geq 0$, επειδή:

$$h^2 - h^{+2} = \begin{cases} h^2, & \text{αν } \; h < 0 \\ 0\,, & \text{αν } \; h > 0 \end{cases}$$

Επίσης το B, με τη βοήθεια του Λήμματος $5(iv)$ (βλ.Παράρτημα), γράφεται:

$$\begin{aligned} B &= \int_{\mathbb{R}^{n-k}} \left(\int_{\mathbb{R}^k} h^-\Big(\|x\|^2, \frac{n-k}{\|z\|^2}\Big)\, {}^t(x-\theta) x N_k(\theta, \sigma^2 I_k; dx) \right) N_{n-k}(0, \sigma^2 I_{n-k}; dz) \\ &= \sigma^2 \int_{\mathbb{R}^{n-k}} \left(\int_{\mathbb{R}^k} h^-\Big(\sigma^2\|t\|^2, \frac{n-k}{\|z\|^2}\Big)\, {}^t(t - \frac{\theta}{\sigma}) t N_k\Big(\frac{\theta}{\sigma}, I_k; dt\Big) \right) N_{n-k}(0, \sigma^2 I_{n-k}; dz) \\ &= \sigma^2 \int_{\mathbb{R}^{n-k}} \left(\int_{\mathbb{N}_0} \left(\int_{\mathbb{R}_+} h^-\Big(\sigma^2 w, \frac{n-k}{\|z\|^2}\Big)(w - 2m)\mathcal{X}^2_{k+2m}(dw) \right) \pi\Big(\frac{\|\theta\|^2}{2\sigma^2}; dm\Big) \right) N_{n-k}(0, \sigma^2 I_{n-k}; dz) \end{aligned}$$

Όμως, εξ υποθέσεως, η συνάρτηση $w \rightsquigarrow h^-(\sigma^2 w, (n-k)/\|z\|^2)$ ανήκει στην κλάση $\mathcal{C}$ και επομένως, $\forall m \in \mathbb{N}_0$,

$$\int_{\mathbb{R}_+} h^-\Big(\sigma^2 w, \frac{n-k}{\|z\|^2}\Big)(w - 2m)\mathcal{X}^2_{k+2m}(dw) \geq 0$$

Λοιπόν και το $B \geq 0$. Κατά συνέπεια έχουμε ότι:

$$\forall (\theta, \sigma) \quad , \quad R_\phi(\theta, \sigma) \geq R_{\phi_{h^+}}(\theta, \sigma)$$

το οποίο ολοκληρώνει την απόδειξη.

□

4.6.2 "Κόψιμο" εκ των άνω

Η "προσέγγιση", η οποία πραγματοποιείται με την συνάρτηση συρρικνώσεως h, φαίνεται να διαψεύδεται όταν η $h(\overline{v_\Theta^{-1}}(\phi_0(y)), 1/s^2(y)) \geq 1$. Στην περίπτωση αυτή οι $\phi(y)$ και $\phi_0(y)$ βρίσκονται από το ένα και το άλλο μέρος του πόλου 0.

Αυτό υποδεικνύει την αντικατάσταση της h από την:

$$h_{\square} = inf(h, 1) = 1 - (1-h)^+$$

Η Πρόταση που ακολουθεί δικαιολογεί αυτή τη μέθοδο για όλους τους εκτιμητές του τύπου (4.6)

Πρόταση 4.6.7

Για κάθε εκτιμητή ϕ_h του τύπου (4.6), ο εκτιμητής που λαμβάνεται "κόβοντας" εκ των άνω την h στο 1, $\phi_{h_\square}$, είναι ομοιομόρφως καλύτερος, δηλαδή,

$$\forall (\theta, \sigma) \quad , \quad R_{\phi_h}(\theta, \sigma) \geq R_{\phi_{h_\square}}(\theta, \sigma)$$

Απόδειξη

Εργαζόμενοι όπως στην απόδειξη της Πρότ.4.6.6, έχουμε:

$$
\begin{aligned}
R_{\phi_h}(\theta,\sigma) - R_{\phi_{h_\square}}(\theta,\sigma) &= \frac{1}{\sigma^2}\int_E \Big\{ \overline{v_\Theta^{-1}}\Big(\Big(1-h(\overline{v_\Theta^{-1}}(\phi_0(y)),1/s^2(y))\Big)\phi_0(y)-\theta\Big) \\
&\qquad - \overline{v_\Theta^{-1}}\Big(\Big(1-h(\overline{v_\Theta^{-1}}(\phi_0(y)),1/s^2(y))\Big)^+\phi_0(y)-\theta\Big)\Big\} N_E(\theta,\sigma^2 v;dy) \\
&= \frac{1}{\sigma^2}(A+2B)
\end{aligned}
$$

όπου

$$
A = \int_E \Big\{ \Big[1-h\Big(\overline{v_\Theta^{-1}}(\phi_0(y)),\frac{1}{s^2(y)}\Big)\Big]^2 - \Big[1-h\Big(\overline{v_\Theta^{-1}}(\phi_0(y)),\frac{1}{s^2(y)}\Big)\Big]^{+2}\Big\}\overline{v_\Theta^{-1}}(\phi_0(y))N_E(\theta,\sigma^2 v;dy)
$$

και

$$
B = \int_E \Big[1-h\Big(\overline{v_\Theta^{-1}}(\phi_0(y)),\frac{1}{s^2(y)}\Big)\Big]^{-}\ {}^t\phi_0(y)\theta N_E(\theta,\sigma^2 v;dy)
$$

Προφανώς $A \geq 0$, επειδή

$$
(1-h)^2-(1-h)^{+2} = \begin{cases} (1-h)^2, & \text{αν } (1-h)<0 \\ 0 \quad , & \text{αν } (1-h)>0 \end{cases}
$$

και χρησιμοποιώντας το Λήμμα 5(iii) (βλ.Παράρτημα),

$$
\begin{aligned}
B &= \int_{\Theta\times\Theta^\perp}\Big[1-h\Big(\overline{v_\Theta^{-1}}(x),\frac{n-k}{\overline{v_{\Theta^\perp}^{-1}}(z)}\Big)\Big]^{-}\ {}^t x\theta N_\Theta(\theta,\sigma^2 v_\Theta;dx)N_{\Theta^\perp}(0,\sigma^2 v_{\Theta^\perp};dz) \\
&= \Big[\int_{\mathbb{R}^{n-k}}\Big(\int_{\mathbb{R}^k}\Big[1-h\Big(\|x\|^2,\frac{n-k}{\|z\|^2}\Big)\Big]^{-}{}^t x N_k(\theta,\sigma^2 I_k;dx)\Big)N_{n-k}(0,\sigma^2 I_k;dz)\Big]\theta \\
&= \sigma\Big[\int_{\mathbb{R}^{n-k}}\Big(\int_{\mathbb{R}^k}\Big[1-h\Big(\sigma^2\|t\|^2,\frac{n-k}{\|z\|^2}\Big)\Big]^{-}{}^t t N_k\Big(\frac{\theta}{\sigma},I_k;dt\Big)\Big)N_{n-k}(0,\sigma^2 I_k;dz)\Big]\theta \\
&= \sigma^t\theta\Big[\int_{\mathbb{R}^{n-k}}\Big(\int_{\mathbb{N}_0}\Big(\int_{\mathbb{R}_+}\Big[1-h\Big(\sigma^2 w,\frac{n-k}{\|z\|^2}\Big)\Big]^{-}\mathcal{X}^2_{k+2+2m}(dw)\Big)\pi\Big(\frac{\|\theta\|^2}{2\sigma^2};dm\Big)\Big)N_{n-k}(0,\sigma^2 I_k;dz)\Big]\theta \\
&= \sigma\|\theta\|^2\int_{\mathbb{R}^{n-k}}\Big(\int_{\mathbb{N}_0}\Big(\int_{\mathbb{R}_+}\Big[1-h\Big(\sigma^2 w,\frac{n-k}{\|z\|^2}\Big)\Big]^{-}\mathcal{X}^2_{k+2+2m}(dw)\Big)\pi\Big(\frac{\|\theta\|^2}{2\sigma^2};dm\Big)\Big)N_{n-k}(0,\sigma^2 I_k;dz) \\
&\geq 0
\end{aligned}
$$

το οποίο ολοκληρώνει την απόδειξη.

$\square$

4.6.3 Διπλό ‘‘Κόψιμο’’

Αν ο εκτιμητής ϕ_h είναι του τύπου (4.6) με την h να λαμβάνει συγχρόνως τιμές αυστηρά αρνητικές και τιμές αυστηρά μεγαλύτερες του 1 και αν επιπλέον η h ικανοποιεί τις υποθέσεις της Πρότασης 4.6.6 μπορούμε βεβαίως να τον βελτιώσουμε χρησιμοποιώντας τον $\phi_{h_\square^+}$ με:

$$
\begin{aligned}
h_\square^+ &= \begin{cases} 1, & \text{αν} \quad h>1 \\ h, & \text{αν} \quad 0\leq h\leq 1 \\ 0, & \text{αν} \quad h<0 \end{cases} \\
&= \inf(1,h^+)
\end{aligned}
$$

4.7 ΜΕΡΙΚΗ ΠΕΡΙΠΤΩΣΗ: ΕΚΤΙΜΗΤΗΣ ΤΩΝ JAMES – STEIN "ΘΕΤΙΚΟΥ ΜΕΡΟΥΣ"

Μία ενδιαφέρουσα υποκλάση της (4.6) είναι η κλάση των εκτιμητών των *James – Stein* της μορφής:

$$\phi(y) = \left(1 - \alpha \frac{\overline{v_{\Theta^\perp}^{-1}}(y - \phi_0(y))}{\overline{v_\Theta^{-1}}(\phi_0(y))}\right)\phi_0(y)\ , \quad \alpha \in \mathbb{R}_+$$

στην οποία υποκλάση ο "βέλτιστος"(*optimal*) εκτιμητής, υπό τη συνήθη τετραγωνική απώλεια, $\sigma^{-2}\overline{v_\Theta^{-1}}(\widehat{\theta} - \theta)$, είναι εκείνος των *James – Stein*

$$\phi_{JS}(y) = \left(1 - \frac{k-2}{n-k+2}\ \frac{\overline{v_{\Theta^\perp}^{-1}}(y - \phi_0(y))}{\overline{v_\Theta^{-1}}(\phi_0(y))}\right)\phi_0(y) \tag{4.9}$$

Ο εκτιμητής αυτός βελτιώνεται με "κόψιμο" εκ των άνω (βλ.4.6.2) από τον εκτιμητή *James– Stein* "θετικού μέρους":

$$\phi_{JS}^+(y) = \left(1 - \frac{k-2}{n-k+2}\ \frac{\overline{v_{\Theta^\perp}^{-1}}(y - \phi_0(y))}{\overline{v_\Theta^{-1}}(\phi_0(y))}\right)^+ \phi_0(y) \tag{4.9$'$}$$

Αν και αυτός ο εκτιμητής δεν είναι παραδεκτός, κανένας εκτιμητής ομοιμόρφως καλύτερος από αυτόν δεν είναι γνωστός. Οι *Bock*[1987] και *Brown*[1987] αποδεικνύουν ότι η συνήθης τεχνική της αμερόληπτης εκτίμησης του κινδύνου δεν είναι χρήσιμη σε αυτήν την περίπτωση.

4.7.1 Κίνδυνος του εκτιμητή των James – Stein ($\phi_{\mathbf{JS}}$)

Ο κίνδυνος του εκτιμητή ϕ_{JS} είναι:

$$\begin{aligned}
&R_{\phi_{JS}}(\theta, \sigma)\\
&= \sigma^{-2}\int_E \overline{v_\Theta^{-1}}\Big(\phi_0(y) - \theta - \frac{k-2}{n-k+2}\ \frac{\overline{v_{\Theta^\perp}^{-1}}(y - \phi_0(y))}{\overline{v_\Theta^{-1}}(\phi_0(y))}\phi_0(y)\Big) N_E(\theta, \sigma^2 v; dy)\\
&= \sigma^{-2}\int_{\Theta\times\Theta^\perp} \overline{v_\Theta^{-1}}\Big(x - \theta - \frac{k-2}{n-k+2}\ \frac{\overline{v_{\Theta^\perp}^{-1}}(z)}{\overline{v_\Theta^{-1}}(x)}x\Big) N_\Theta(\theta, \sigma^2 v_\Theta; dx) N_{\Theta^\perp}(0, \sigma^2 v_{\Theta^\perp}; dz)\\
&= \sigma^{-2}\int_{\mathbb{R}^{n-k}}\int_{\mathbb{R}^k} \Big\|x - \theta - \frac{k-2}{n-k+2}\ \frac{\|z\|^2}{\|x\|^2}x\Big\| N_k(\theta, \sigma^2 I_k; dx) N_{n-k}(0, \sigma^2 I_{n-k}; dz)\\
&= \sigma^{-2}\int_{\mathbb{R}^{n-k}}\int_{\mathbb{R}^k} \Big\{\|x-\theta\|^2 + \frac{(k-2)^2}{(n-k+2)^2}\ \frac{\|z\|^4}{\|x\|^2} - 2\frac{k-2}{n-k+2}\ \frac{\|z\|^2}{\|x\|^2}{}^t(x-\theta)x\Big\} N_k(\theta, \sigma^2 I_k; dx)\\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad N_{n-k}(0, \sigma^2 I_{n-k}; dz)\\
&= A + B + C
\end{aligned}$$

όπου, υπό τη χρήση του Παραρτήματος (Λήμμα 1(3), Πόρισμα 1, Λήμμα 5(iv)), αντιστοίχως επί των ολοκληρωμάτων A, B και C, λαμβάνουμε:

$$A \ = \ \sigma^{-2}\int_{\mathbb{R}^{n-k}}\int_{\mathbb{R}^k} \|x-\theta\|^2 N_k(\theta, \sigma^2 I_k; dx) N_{n-k}(0, \sigma^2 I_{n-k}; dz)$$

$$
\begin{aligned}
&= \int_{\mathbb{R}^k}\int_{\mathbb{R}^{n-k}} \Big\| \frac{x-\theta}{\sigma} \Big\|^2 N_k(\theta,\sigma^2 I_k;dx) N_{n-k}(0,\sigma^2 I_{n-k};dz) \\
&= \int_{\mathbb{R}^{n-k}}\int_{\mathbb{R}^k} \|t\|^2 N_k(0,I_k;dt) N_{n-k}(0,\sigma^2 I_{n-k};dz) \\
&= \int_{\mathbb{R}^{n-k}}\int_{\mathbb{R}_+} u\mathcal{X}_k^2(du) N_{n-k}(0,\sigma^2 I_{n-k};dz) \\
&= \int_{\mathbb{R}^{n-k}} kN_{n-k}(0,\sigma^2 I_{n-k};dz) = k
\end{aligned}
$$

$$
\begin{aligned}
B &= \sigma^{-2}\frac{(k-2)^2}{(n-k+2)^2}\int_{\mathbb{R}^{n-k}}\int_{\mathbb{R}^k} \frac{\|z\|^4}{\|x\|^2} N_k(\theta,\sigma^2 I_k;dx) N_{n-k}(0,\sigma^2 I_{n-k};dz) \\
&= \sigma^{-2}\frac{(k-2)^2}{(n-k+2)^2}\int_{\mathbb{R}^{n-k}}\int_{\mathbb{R}^k} \sigma^2\frac{\|z/\sigma\|^4}{\|x/\sigma\|^2} N_k(\theta,\sigma^2 I_k;dx) N_{n-k}(0,\sigma^2 I_{n-k};dz) \\
&= \frac{(k-2)^2}{(n-k+2)^2}\int_{\mathbb{R}^{n-k}}\int_{\mathbb{R}^k} \frac{\|s\|^4}{\|t\|^2} N_k\Big(\frac{\theta}{\sigma},I_k;dt\Big) N_{n-k}(0,I_{n-k};ds) \\
&= \frac{(k-2)^2}{(n-k+2)^2}\int_{\mathbb{R}_+}\Big(\int_{\mathbb{R}_+} \frac{v^2}{u}\mathcal{X}_k^2\Big(\frac{\|\theta\|^2}{\sigma^2};du\Big)\Big)\mathcal{X}_{n-k}^2(dv) \\
&= \frac{(k-2)^2}{(n-k+2)^2}\Big(\int_{\mathbb{R}_+} v^2\mathcal{X}_{n-k}^2(dv)\Big)\Big(\int_{\mathbb{R}_+} \frac{1}{u}\mathcal{X}_k^2\Big(\frac{\|\theta\|^2}{\sigma^2};du\Big)\Big) \\
&= \frac{(k-2)^2}{(n-k+2)^2}\;(n-k)(n-k+2)\;\frac{1}{2}\int_{\mathbb{N}_0} \frac{\Gamma(\frac{k}{2}+m-1)}{\Gamma(\frac{k}{2}+m)}\;\pi\Big(\frac{\|\theta\|^2}{2\sigma^2};dm\Big) \\
&= \frac{(k-2)^2(n-k)}{(n-k+2)}\;\frac{1}{2}\int_{\mathbb{N}_0} \frac{\Gamma(\frac{k}{2}+m-1)}{(\frac{k}{2}+m-1)\Gamma(\frac{k}{2}+m-1)}\;\pi\Big(\frac{\|\theta\|^2}{2\sigma^2};dm\Big) \\
&= \frac{(k-2)^2(n-k)}{(n-k+2)}\int_{\mathbb{N}_0} \frac{1}{k+2m-2}\;\pi\Big(\frac{\|\theta\|^2}{2\sigma^2};dm\Big)
\end{aligned}
$$

και

$$
\begin{aligned}
C &= -\frac{2}{\sigma^2}\;\frac{k-2}{n-k+2}\int_{\mathbb{R}^k} \frac{1}{\|x\|^2}\;{}^t(x-\theta)x\Big(\int_{\mathbb{R}^{n-k}} \|z\|^2 N_{n-k}(0,\sigma^2 I_{n-k};dz)\Big) N_k(\theta,\sigma^2 I_k;dx) \\
&= -2\;\frac{k-2}{n-k+2}\Big(\int_{\mathbb{R}^k} \frac{1}{\|x/\sigma\|^2}\;{}^t\Big(\frac{x}{\sigma}-\frac{\theta}{\sigma}\Big)\frac{x}{\sigma} N_k(\theta,\sigma^2 I_k;dx)\Big)\Big(\int_{\mathbb{R}^{n-k}} \|z/\sigma\|^2 N_{n-k}(0,\sigma^2 I_{n-k};dz)\Big) \\
&= -2\;\frac{k-2}{n-k+2}\Big(\int_{\mathbb{R}^k} \frac{1}{\|t\|^2}\;{}^t\Big(t-\frac{\theta}{\sigma}\Big)tN_k\Big(\frac{\theta}{\sigma},I_k;dt\Big)\Big)\Big(\int_{\mathbb{R}^{n-k}} \|s\|^2 N_{n-k}(0,I_{n-k};ds)\Big) \\
&= -2\;\frac{k-2}{n-k+2}\Big(\int_{\mathbb{N}_0}\Big(\int_{\mathbb{R}_+} \frac{1}{u}(u-2m)\mathcal{X}_{k+2m}^2(du)\Big)\pi\Big(\frac{\|\theta\|^2}{2\sigma^2};dm\Big)\Big)\Big(\int_{\mathbb{R}_+} v\mathcal{X}_{n-k}^2(dv)\Big)
\end{aligned}
$$

$$
\begin{aligned}
&= -2\,\frac{k-2}{n-k+2}\,(n-k)\int_{\mathbb{N}_0}\Big(\int_{\mathbb{R}_+}\Big(1-\frac{2m}{u}\Big)\mathcal{X}^2_{k+2m}(du)\Big)\pi\Big(\frac{\|\theta\|^2}{2\sigma^2};dm\Big)\\
&= -2\,\frac{(k-2)(n-k)}{n-k+2}+2\,\frac{(k-2)(n-k)}{n-k+2}\int_{\mathbb{N}_0}2m\Big(\int_{\mathbb{R}_+}\frac{1}{u}\mathcal{X}^2_{k+2m}(du)\Big)\pi\Big(\frac{\|\theta\|^2}{2\sigma^2};dm\Big)\\
&= -2\,\frac{(k-2)(n-k)}{n-k+2}+2\,\frac{(k-2)(n-k)}{n-k+2}\int_{\mathbb{N}_0}2m\Big(\frac{1}{2}\,\frac{\Gamma(\frac{k}{2}+m-1)}{\Gamma(\frac{k}{2}+m)}\Big)\pi\Big(\frac{\|\theta\|^2}{2\sigma^2};dm\Big)\\
&= -2\,\frac{(k-2)(n-k)}{n-k+2}+2\,\frac{(k-2)(n-k)}{n-k+2}\int_{\mathbb{N}_0}\frac{2m}{k+2m-2}\pi\Big(\frac{\|\theta\|^2}{2\sigma^2};dm\Big)\\
&= -2\,\frac{(k-2)(n-k)}{n-k+2}+2\,\frac{(k-2)(n-k)}{n-k+2}\sum_{m=1}^{+\infty}\frac{2m}{k+2(m-1)}\,\frac{1}{m!}exp\Big\{\frac{-\|\theta\|^2}{2\sigma^2}\Big\}\Big(\frac{\|\theta\|^2}{2\sigma^2}\Big)^m\\
&= -2\,\frac{(k-2)(n-k)}{n-k+2}+2\,\frac{(k-2)(n-k)}{n-k+2}\sum_{l=0}^{+\infty}\frac{1}{k+2l}\,\frac{\|\theta\|^2}{\sigma^2}\frac{1}{l!}exp\Big\{\frac{-\|\theta\|^2}{2\sigma^2}\Big\}\Big(\frac{\|\theta\|^2}{2\sigma^2}\Big)^l\\
&= -2\,\frac{(k-2)(n-k)}{n-k+2}+2\,\frac{(k-2)(n-k)}{n-k+2}\,\frac{\|\theta\|^2}{\sigma^2}\int_{\mathbb{N}_0}\frac{1}{k+2l}\,\pi\Big(\frac{\|\theta\|^2}{2\sigma^2};dl\Big)
\end{aligned}
$$

Συνεπώς,

$$
\begin{aligned}
R_{\phi_{JS}}(\theta,\sigma) &= k+\frac{(k-2)^2(n-k)}{(n-k+2)}\int_{\mathbb{N}_0}\frac{1}{k+2m-2}\,\pi\Big(\frac{\|\theta\|^2}{2\sigma^2};dm\Big)\\
&\quad -2\,\frac{(k-2)(n-k)}{n-k+2}+2\,\frac{(k-2)(n-k)}{n-k+2}\,\frac{\|\theta\|^2}{\sigma^2}\int_{\mathbb{N}_0}\frac{1}{k+2m}\,\pi\Big(\frac{\|\theta\|^2}{2\sigma^2};dm\Big) \qquad (4.10)
\end{aligned}
$$

Όμως, θέτοντας $\lambda=\frac{\|\theta\|^2}{2\sigma^2}$ και χρησιμοποιώντας τη σχέση:

$$\int_0^\lambda x^{m+\frac{k}{2}-1}dx=\frac{\lambda^{m+\frac{k}{2}}}{m+\frac{k}{2}}$$

έχουμε:

$$
\begin{aligned}
\int_{\mathbb{N}_0}\frac{1}{\frac{k}{2}+m}\,\pi(\lambda;dm) &= \sum_{m=0}^{+\infty}\frac{e^{-\lambda}\lambda^m}{m!}\,\frac{1}{\frac{k}{2}+m}=\frac{1}{\lambda^{k/2}}\sum_{m=0}^{+\infty}\frac{e^{-\lambda}\lambda^{m+\frac{k}{2}}}{m!(\frac{k}{2}+m)}\\
&= \frac{1}{\lambda^{k/2}}\sum_{m=0}^{+\infty}\frac{e^{-\lambda}}{m!}\int_0^\lambda x^{m+\frac{k}{2}-1}dx=\frac{e^{-\lambda}}{\lambda^{k/2}}\int_0^\lambda x^{\frac{k}{2}-1}\Big(\sum_{m=0}^{+\infty}\frac{x^m}{m!}\Big)dx\\
&= \frac{e^{-\lambda}}{\lambda^{k/2}}\int_0^\lambda x^{\frac{k}{2}-1}e^x dx
\end{aligned}
$$

Ομοίως προκύπτει ότι:

$$\int_{\mathbb{N}_0}\frac{1}{\frac{k}{2}+m-1}\,\pi(\lambda;dm)=\frac{e^{-\lambda}}{\lambda^{\frac{k}{2}-1}}\int_0^\lambda x^{\frac{k}{2}-2}e^x dx$$

Άρα, τελικά, η σχέση (4.10) γράφεται:

$$\begin{aligned} R_{\phi_{JS}}(\theta,\sigma) &= k + \frac{(k-2)^2(n-k)}{(n-k+2)} \frac{e^{-\lambda}}{2\lambda^{\frac{k}{2}-1}} \int_0^\lambda x^{\frac{k}{2}-2} e^x dx \\ & \quad -2\, \frac{(k-2)(n-k)}{n-k+2} + 2\, \frac{(k-2)(n-k)}{n-k+2} \lambda \frac{e^{-\lambda}}{\lambda^{k/2}} \int_0^\lambda x^{\frac{k}{2}-1} e^x dx \end{aligned} \tag{4.11}$$

με $\lambda = \dfrac{\|\theta\|^2}{2\sigma^2}$

4.7.2 Κέρδος του εκτιμητή των James – Stein "θετικού μέρους" ($\phi^+_{\mathbf{JS}}$)

Ο εκτιμητής ϕ^+_{JS} είναι προφανώς ένας εκτιμητής της μορφής (4.6) :

$$\phi^+_{JS}(y) = \left(1 - h\Big(\overline{v_\Theta^{-1}}(\phi_0(y)), \frac{1}{s^2(y)}\Big)\right)^+ \phi_0(y)$$

όπου

$$h\Big(\overline{v_\Theta^{-1}}(\phi_0(y)), \frac{1}{s^2(y)}\Big) = \frac{\alpha_0 s^2(y)}{\overline{v_\Theta^{-1}}(\phi_0(y))}\, I\Big(\frac{\alpha_0 s^2(y)}{\overline{v_\Theta^{-1}}(\phi_0(y))} < 1\Big) + I\Big(\frac{\alpha_0 s^2(y)}{\overline{v_\Theta^{-1}}(\phi_0(y))} > 1\Big)$$

με $\quad \alpha_0 = \dfrac{(k-2)(n-k)}{n-k+2}$

$$I\Big(\frac{\alpha_0 s^2(y)}{\overline{v_\Theta^{-1}}(\phi_0(y))} < 1\Big) = \begin{cases} 1 & \text{αν} \quad \frac{\alpha_0 s^2(y)}{\overline{v_\Theta^{-1}}(\phi_0(y))} < 1 \\ 0 & \qquad \text{αλλού} \end{cases}$$

και

$$I\Big(\frac{\alpha_0 s^2(y)}{\overline{v_\Theta^{-1}}(\phi_0(y))} > 1\Big) = 1 - I\Big(\frac{\alpha_0 s^2(y)}{\overline{v_\Theta^{-1}}(\phi_0(y))} < 1\Big)$$

Τότε η διαφορά των κινδύνων μεταξύ των εκτιμητών ϕ_0 και ϕ^+_{JS} είναι:

$$\begin{aligned} R_{\phi_0}(\theta,\sigma) - R_{\phi^+_{JS}}(\theta,\sigma) &= \sigma^{-2}\mathbb{E}_{\theta,\sigma}\Big[\overline{v_\Theta^{-1}}(\phi_0(y)-\theta)\Big] - \sigma^{-2}\mathbb{E}_{\theta,\sigma}\Big[\overline{v_\Theta^{-1}}(\phi^+_{JS}(y)-\theta)\Big] \\ &= 2\sigma^{-2}\mathbb{E}_{\theta,\sigma}\Big[h\Big(\overline{v_\Theta^{-1}}(\phi_0(y)), \frac{1}{s^2(y)}\Big) v_\Theta^{-1}\Big(\phi_0(y), \phi_0(y)-\theta\Big)\Big] \\ & \quad -\sigma^{-2}\mathbb{E}_{\theta,\sigma}\Big[h^2\Big(\overline{v_\Theta^{-1}}(\phi_0(y)), \frac{1}{s^2(y)}\Big)\overline{v_\Theta^{-1}}(\phi_0(y))\Big] \end{aligned}$$

Όμως, βάσει του Λήμματος 5(iv)(βλ.Παράρτημα),

$$\begin{aligned} & \sigma^{-2}\mathbb{E}_{\theta,\sigma}\Big[h\Big(\overline{v_\Theta^{-1}}(\phi_0(y)), \frac{1}{s^2(y)}\Big) v_\Theta^{-1}\Big(\phi_0(y), \phi_0(y)-\theta\Big)\Big] \\ &= \sigma^{-2}\int_{\Theta\times\Theta^\perp} h\Big(\overline{v_\Theta^{-1}}(x), \frac{n-k}{\overline{v_{\Theta^\perp}^{-1}}(z)}\Big) v_\Theta^{-1}(x, x-\theta) N_\Theta(\theta, \sigma^2 v_\Theta; dx) N_{\Theta^\perp}(0, \sigma^2 v_{\Theta^\perp}; dz) \\ &= \sigma^{-2}\int_{\mathbb{R}^k\times\mathbb{R}^{n-k}} h\Big(\|x\|^2, \frac{n-k}{\|z\|^2}\Big)^t x(x-\theta) N_k(\theta, \sigma^2 I_k; dx) N_{n-k}(0, \sigma^2 I_{n-k}; dz) \\ &= \int_{\mathbb{R}_+}\Big[\int_{\mathbb{N}_0}\Big(\int_{\mathbb{R}_+} h\Big(\sigma^2 u, \frac{n-k}{\sigma^2 s}\Big)(u-2m)\mathcal{X}^2_{k+2m}(du)\Big)\pi\Big(\frac{\|\theta\|^2}{2\sigma^2}; dm\Big)\Big]\mathcal{X}^2_{n-k}(ds) \end{aligned} \tag{4.12}$$

και βάσει της έκφρασης (3) του Λήμματος 1(βλ.Παράρτημα) έχουμε ότι:

$$
\begin{aligned}
&\sigma^{-2}\mathbb{E}_{\theta,\sigma}\Big[h^2\Big(\overline{v_\Theta^{-1}}(\phi_0(y)), \frac{1}{s^2(y)}\Big)\overline{v_\Theta^{-1}}(\phi_0(y))\Big] \\
&= \sigma^{-2}\int_{\Theta\times\Theta^\perp} h^2\Big(\overline{v_\Theta^{-1}}(x), \frac{n-k}{\overline{v_{\Theta^\perp}^{-1}}(z)}\Big)\overline{v_\Theta^{-1}}(x)N_\Theta(\theta,\sigma^2 v_\Theta; dx)N_{\Theta^\perp}(0,\sigma^2 v_{\Theta^\perp}; dz) \\
&= \sigma^{-2}\int_{\mathbb{R}^k\times\mathbb{R}^{n-k}} h^2\Big(\|x\|^2, \frac{n-k}{\|z\|^2}\Big)\|x\|^2 N_k(\theta,\sigma^2 I_k; dx)N_{n-k}(0,\sigma^2 I_{n-k}; dz) \\
&= \int_{\mathbb{R}_+}\Big(\int_{\mathbb{R}_+} h^2\Big(\sigma^2 u, \frac{n-k}{\sigma^2 s}\Big)u\mathcal{X}_k^2\Big(\frac{\|\theta\|^2}{\sigma^2}; du\Big)\Big)\mathcal{X}_{n-k}^2(ds) \\
&= \int_{\mathbb{R}_+}\Big[\int_{\mathbb{N}_0}\Big(\int_{\mathbb{R}_+} h^2\Big(\sigma^2 u, \frac{n-k}{\sigma^2 s}\Big)u\mathcal{X}_{k+2m}^2(du)\Big)\pi\Big(\frac{\|\theta\|^2}{2\sigma^2}; dm\Big)\Big]\mathcal{X}_{n-k}^2(ds) \qquad (4.13)
\end{aligned}
$$

με $\quad h\Big(\sigma^2 u, \frac{n-k}{\sigma^2 s}\Big) = \alpha_0\frac{s}{(n-k)u}\, I\Big(\frac{\alpha_0 s}{(n-k)u} < 1\Big) + I\Big(\frac{\alpha_0 s}{(n-k)u} > 1\Big)$

και $\quad h^2\Big(\sigma^2 u, \frac{n-k}{\sigma^2 s}\Big) = \alpha_0^2\frac{s^2}{(n-k)^2u^2}\, I\Big(\frac{\alpha_0 s}{(n-k)u} < 1\Big) + I\Big(\frac{\alpha_0 s}{(n-k)u} > 1\Big)$

Συνεπώς, λόγω των (4.12) και (4.13), προκύπτει ότι:

$$
\begin{aligned}
&R_{\phi_0}(\theta,\sigma) - R_{\phi_{JS}^+}(\theta,\sigma) \\
&= 2\int_{\mathbb{R}_+\times\mathbb{N}_0\times\mathbb{R}_+}\Big\{\alpha_0\frac{s}{(n-k)u}\, I\Big(\frac{\alpha_0 s}{(n-k)u} < 1\Big) + I\Big(\frac{\alpha_0 s}{(n-k)u} > 1\Big)\Big\}(u-2m)\mathcal{X}_{k+2m}^2(du) \\
&\qquad\qquad \pi\Big(\frac{\|\theta\|^2}{2\sigma^2}; dm\Big)\mathcal{X}_{n-k}^2(ds) \\
&- \int_{\mathbb{R}_+\times\mathbb{N}_0\times\mathbb{R}_+}\Big\{\alpha_0^2\frac{s^2}{(n-k)^2u^2}\, I\Big(\frac{\alpha_0 s}{(n-k)u} < 1\Big) + I\Big(\frac{\alpha_0 s}{(n-k)u} > 1\Big)\Big\}u\mathcal{X}_{k+2m}^2(du) \\
&\qquad\qquad \pi\Big(\frac{\|\theta\|^2}{2\sigma^2}; dm\Big)\mathcal{X}_{n-k}^2(ds) \qquad (4.14)
\end{aligned}
$$

Θέτοντας, σε ότι ακολουθεί, $\beta = \frac{\alpha_0 s}{n-k}$ και $\gamma^{-1} = \frac{\alpha_0}{(n-k)u}$, υπό παραγοντική ολοκλήρωση λαμβάνουμε:

$$
\begin{aligned}
&\int_{\mathbb{R}_+}\Big\{\alpha_0\frac{s}{(n-k)u}\, I\Big(\frac{\alpha_0 s}{(n-k)u} < 1\Big) + I\Big(\frac{\alpha_0 s}{(n-k)u} > 1\Big)\Big\}(u-2m)\mathcal{X}_{k+2m}^2(du) \\
&= \int_{\mathbb{R}_+}\Big\{\beta\frac{1}{u}\, I(u>\beta) + I(u<\beta)\Big\}(u-2m)\mathcal{X}_{k+2m}^2(du) \\
&= \beta\int_\beta^{+\infty}\Big(1-\frac{2u}{\beta}\Big)\mathcal{X}_{k+2m}^2(du) + \int_0^\beta (u-2m)\mathcal{X}_{k+2m}^2(du) \\
&= \frac{2\beta}{\Gamma(\frac{k}{2}+m)2^{\frac{k}{2}+m}}\,\beta^{\frac{k}{2}+m-1}e^{-\beta/2} + (k-2)\beta\int_\beta^{+\infty} u^{-1}\mathcal{X}_{k+2m}^2(du)
\end{aligned}
$$

$$-\frac{2}{\Gamma(\frac{k}{2}+m)2^{\frac{k}{2}+m}}\ \beta^{\frac{k}{2}+m}e^{-\beta/2}+k\int_0^{\beta}\mathcal{X}^2_{k+2m}(du)$$

$$=(k-2)\beta\int_{\beta}^{+\infty}u^{-1}\mathcal{X}^2_{k+2m}(du)+k\int_0^{\beta}\mathcal{X}^2_{k+2m}(du) \tag{4.15}$$

και

$$\int_{\mathbb{R}_+}\Big\{\alpha_0^2\frac{s^2}{(n-k)^2u^2}\ I\Big(\alpha_0\frac{s}{(n-k)u}<1\Big)+I\Big(\alpha_0\frac{s}{(n-k)u}>1\Big)u\mathcal{X}^2_{n-k}(ds)$$

$$=\frac{\alpha_0}{n-k}\gamma^{-1}\int_0^{\gamma}s^2\mathcal{X}^2_{n-k}(ds)+u\int_{\gamma}^{+\infty}\mathcal{X}^2_{n-k}(ds)$$

$$=-2\gamma^{\frac{n-k}{2}}\frac{\alpha_0}{n-k}\ \frac{e^{-\gamma/2}}{\Gamma(\frac{n-k}{2})2^{\frac{n-k}{2}}}+\alpha_0\gamma^{-1}\ \frac{n-k+2}{n-k}\int_0^{\gamma}s\mathcal{X}^2_{n-k}(ds)$$

$$+\frac{2u}{\Gamma(\frac{n-k}{2})2^{\frac{n-k}{2}}}\ \gamma^{\frac{n-k}{2}-1}e^{-\gamma/2}+(n-k-2)u\int_{\gamma}^{+\infty}s^{-1}\mathcal{X}^2_{n-k}(ds)$$

$$=\frac{2\gamma^{\frac{n-k}{2}}e^{-\gamma/2}}{\Gamma(\frac{n-k}{2})2^{\frac{n-k}{2}}}\underbrace{\Big(\gamma^{-1}u-\frac{\alpha_0}{n-k}\Big)}_{=0}+\alpha_0\gamma^{-1}\frac{n-k+2}{n-k}\int_0^{\gamma}s\mathcal{X}^2_{n-k}(ds)$$

$$+(n-k-2)u\int_{\gamma}^{+\infty}s^{-1}\mathcal{X}^2_{n-k}(ds) \tag{4.16}$$

Από τις (4.15) και (4.16), η σχέση (4.14) γράφεται:

$$R_{\phi_0}(\theta,\sigma)-R_{\phi^+_{JS}}(\theta,\sigma)$$

$$=2\int_{\mathbb{R}_+\times\mathbb{R}_+}\Big\{(k-2)\frac{\alpha_0 s}{(n-k)u}\ I\Big(\frac{\alpha_0 s}{(n-k)u}<1\Big)+k\ I\Big(\frac{\alpha_0 s}{(n-k)u}>1\Big)\Big\}\mathcal{X}^2_k\Big(\frac{\|\theta\|^2}{\sigma^2};du\Big)\mathcal{X}^2_{n-k}(ds)$$

$$-\int_{\mathbb{R}_+\times\mathbb{R}_+}\Big\{\alpha_0^2\frac{n-k+2}{(n-k)^2}\ \frac{s}{u}\ I\Big(\frac{\alpha_0 s}{(n-k)u}<1\Big)+(n-k-2)\frac{u}{s}\ I\Big(\frac{\alpha_0 s}{(n-k)u}>1\Big)\Big\}\mathcal{X}^2_k\Big(\frac{\|\theta\|^2}{\sigma^2};du\Big)$$
$$\mathcal{X}^2_{n-k}(ds)$$

Αυτό το κέρδος μπορεί επίσης να γραφεί:

$$R_{\phi_0}(\theta,\sigma)-R_{\phi^+_{JS}}(\theta,\sigma)$$

$$=k-\frac{\|\theta\|^2}{\sigma^2}+\int_{\mathbb{R}_+\times\mathbb{R}_+}\Big\{\frac{(k-2)^2}{n-k+2}\ \frac{s}{u}-2k+(n-k-2)\frac{u}{s}\Big\}I\Big(\frac{\alpha_0 s}{(n-k)u}>1\Big)\mathcal{X}^2_k\Big(\frac{\|\theta\|^2}{\sigma^2};du\Big)\mathcal{X}^2_{n-k}(ds) \tag{4.17}$$

Πράγματι, επειδή:

$$2\Big\{(k-2)\frac{\alpha_0 s}{(n-k)u}\ I\Big(\frac{\alpha_0 s}{(n-k)u}<1\Big)+k\ I\Big(\frac{\alpha_0 s}{(n-k)u}>1\Big)\Big\}$$

$$=2\Big\{(k-2)\frac{\alpha_0 s}{(n-k)u}\ I\Big(\frac{\alpha_0 s}{(n-k)u}<1\Big)+k-k\ I\Big(\frac{\alpha_0 s}{(n-k)u}<1\Big)\Big\}$$

$$=2k+2\Big((k-2)\frac{\alpha_0 s}{(n-k)u}-k\Big)I\Big(\frac{\alpha_0 s}{(n-k)u}<1\Big)$$

και

$$\alpha_0^2\frac{n-k+2}{(n-k)^2}\ \frac{s}{u}\ I\Big(\frac{\alpha_0 s}{(n-k)u}<1\Big)+(n-k-2)\frac{u}{s}\ I\Big(\frac{\alpha_0 s}{(n-k)u}>1\Big)$$
$$=(n-k-2)\frac{u}{s}+\Big(\alpha_0^2\frac{n-k+2}{(n-k)^2}\ \frac{s}{u}-(n-k-2)\frac{u}{s}\Big)I\Big(\frac{\alpha_0 s}{(n-k)u}<1\Big)$$

έχουμε:

$$R_{\phi_0}(\theta,\sigma)-R_{\phi_{JS}^+}(\theta,\sigma)\ =\ 2k-(n-k-2)\int_{\mathbb{R}_+\times\mathbb{R}_+}\frac{u}{s}\ \mathcal{X}_k^2\Big(\frac{\|\theta\|^2}{\sigma^2};du\Big)\mathcal{X}_{n-k}^2(ds)$$
$$+\int_{\mathbb{R}_+\times\mathbb{R}_+}\Big\{2\Big((k-2)\frac{\alpha_0 s}{(n-k)u}-k\Big)-\Big(\alpha_0^2\frac{n-k+2}{(n-k)^2}\ \frac{s}{u}-(n-k-2)\frac{u}{s}\Big)\Big\}I\Big(\frac{\alpha_0 s}{(n-k)u}<1\Big)$$
$$\mathcal{X}_k^2\Big(\frac{\|\theta\|^2}{\sigma^2};du\Big)\mathcal{X}_{n-k}^2(ds)$$

όπου, από τον ορισμό του α_0,

$$2\Big((k-2)\frac{\alpha_0 s}{(n-k)u}-k\Big)-\Big(\alpha_0^2\frac{n-k+2}{(n-k)^2}\ \frac{s}{u}-(n-k-2)\ \frac{u}{s}\Big)$$
$$=\frac{s}{u}\Big(2(k-2)\frac{\alpha_0}{n-k}-\alpha_0^2\frac{n-k+2}{(n-k)^2}\Big)+(n-k-2)\frac{u}{s}-2k$$
$$=\frac{(k-2)^2}{n-k+2}\ \frac{s}{u}-2k+(n-k-2)\frac{u}{s}$$

και

$$\int_{\mathbb{R}_+\times\mathbb{R}_+}\frac{u}{s}\ \mathcal{X}_k^2\Big(\frac{\|\theta\|^2}{\sigma^2};du\Big)\mathcal{X}_{n-k}^2(ds)=\Big(\int_{\mathbb{R}_+}u\mathcal{X}_k^2\Big(\frac{\|\theta\|^2}{\sigma^2};du\Big)\Big)\Big(\int_{\mathbb{R}_+}s^{-1}\mathcal{X}_{n-k}^2(ds)\Big)$$
$$=\Big(2\int_{\mathbb{N}_0}\frac{\Gamma(\frac{k}{2}+m+1)}{\Gamma(\frac{k}{2}+m)}\ \pi\Big(\frac{\|\theta\|^2}{2\sigma^2};dm\Big)\Big)\Big(2^{-1}\ \frac{\Gamma(\frac{n-k}{2}-1)}{\Gamma(\frac{n-k}{2})}\Big)$$
$$=\Big(\frac{n-k}{2}-1\Big)^{-1}\int_{\mathbb{N}_0}\Big(\frac{k}{2}+m\Big)\ \pi\Big(\frac{\|\theta\|^2}{2\sigma^2};dm\Big)$$
$$=\frac{2}{n-k-2}\Big(\frac{k}{2}+\frac{\|\theta\|^2}{2\sigma^2}\Big)$$
$$=\frac{1}{n-k-2}\Big(k+\frac{\|\theta\|^2}{\sigma^2}\Big)$$

απ᾽ όπου προκύπτει η (4.17).

Η αναλυτική έκφραση του κέρδους εξαρτάται από την φύση του k(δηλ. αν k είναι άρτιος ή περιττός). Στην αναλυτική αυτή έκφραση του κέρδους εμφανίζεται το *Dawson*-ολοκλήρωμα:

$$D(\lambda)=e^{-\lambda^2}\int_0^\lambda e^{t^2}dt$$

το οποίο δίνεται σε πίνακες στους *Abramowitz* και *Stegun*[1964]. Αξίζει να σημειώσουμε ότι η προσέγγιση του e^{t^2} από το $\sum_{i=0}^{100}\frac{t^{2i}}{i!}$ στο παραπάνω ολοκλήρωμα δίνει ακριβώς τα ίδια αποτελέσματα όπως ο πίνακας των *Abramowitz* και *Stegun*.

Λοιπόν, θέτοντας $\lambda = \|\theta\|^2/\sigma^2$, μετά από αρκετά εκτεταμένο λογισμό (βλ.$C.Robert$[1987]([40])), προκύπτει:

$a)$ Αν το k είναι άρτιος, το κέρδος είναι:

$$
\begin{aligned}
& R_{\phi_0}(\theta,\sigma) - R_{\phi^+_{JS}}(\theta,\sigma) \\
&= \frac{(k-2)(n-k)}{n-k+2}\Bigg[\sum_{j=1}^{\frac{k-2}{2}}(-1)^{j+1}\frac{\Gamma(\frac{k}{2})}{\Gamma(\frac{k}{2}-j)}\Big(\frac{2}{\lambda}\Big)^j\Big\{1-\Big(\frac{k-2}{n}\Big)^{-j+\frac{k-2}{2}} e^{-\frac{\lambda}{2}\,\frac{n-k+2}{n}}\Big\}\Bigg] \\
& - e^{-\frac{\lambda}{2}\,\frac{n-k+2}{n}}\Bigg[\sum_{i=1}^{\frac{n-k}{2}}\Big\{(n-k)(n-k+2)-4\,\frac{k(n-k+2)}{k-2}\,i+4i(i-1)\Big\}\Big(\frac{n-k+2}{n}\Big)^{i-2} \\
& \times \frac{\Gamma(\frac{k-2}{2}+i)}{\Gamma(\frac{k}{2})}\,\frac{1}{2i!}\sum_{j=0}^{i-1}\binom{i-1}{j}\frac{\Gamma(\frac{k+2}{2}-j)}{\Gamma(\frac{k+2}{2})}\Big(\frac{k-2}{n}\Big)^j\Big(\frac{\lambda}{2}\Big)^j\Bigg]\Big(\frac{k-2}{n}\Big)^{\frac{k+2}{2}} \qquad (4.18)
\end{aligned}
$$

$b)$ Αν το k είναι περιττός, το κέρδος είναι:

$$
\begin{aligned}
& R_{\phi_0}(\theta,\sigma) - R_{\phi^+_{JS}}(\theta,\sigma) \\
&= \frac{(k-2)(n-k)}{n-k+2}\Bigg[\sum_{j=1}^{\frac{k-3}{2}}(-1)^{j+1}\frac{(k-2)\ldots(k-2j)}{\lambda^j}\Big\{1-\Big(\frac{k-2}{n}\Big)^{-j+\frac{k-2}{2}} e^{-\frac{\lambda}{2}\,\frac{n-k+2}{n}}\Big\} \\
& + \frac{(-1)^{\frac{k-3}{2}}(k-2)!}{(2\lambda)^{\frac{k-3}{2}}(\frac{k-3}{2})!\sqrt{\lambda/2}}\Big\{D(\sqrt{\lambda/2}) - e^{-\frac{\lambda}{2}\,\frac{n-k+2}{n}}D\Big(\Big(\frac{(k-2)\lambda}{2n}\Big)^{1/2}\Big)\Big\}\Bigg] \\
& - e^{-\frac{\lambda}{2}\,\frac{n-k+2}{n}}\Bigg[\sum_{i=1}^{\frac{n-k}{2}}\Big\{(n-k)(n-k+2)-4\,\frac{k(n-k+2)}{k-2}\,i+4i(i-1)\Big\}\Big(\frac{n-k+2}{n}\Big)^{i-2} \\
& \times \frac{\Gamma(\frac{k-2}{2}+i)}{\Gamma(\frac{k}{2})}\,\frac{1}{2i!}\sum_{j=0}^{i-1}\binom{i-1}{j}\frac{\Gamma(\frac{k+2}{2}-j)}{\Gamma(\frac{k+2}{2})}\Big(\frac{k-2}{n}\Big)^j\Big(\frac{\lambda}{2}\Big)^j\Bigg]\Big(\frac{k-2}{n}\Big)^{\frac{k+2}{2}} \qquad (4.19)
\end{aligned}
$$

όπου το $\dfrac{\Gamma(x \pm i)}{\Gamma(x)}$ που εμφανίζεται στους παραπάνω τύπους υπολογίζεται εύκολα από τον τύπο $\Gamma(x+1) = x\Gamma(x)$. Επίσης, η προσέγγιση του e^{t^2}, μέσω του παραπάνω μερικού αθροίσματος, βοηθά στην χρήση των παραπάνω τύπων μέσω Η/Υ.

Παραδείγματα

Αν και οι φόρμουλες (4.18) και (4.19) φαίνονται δυσμεταχείριστες, αυτές μπορούν να περιοριστούν πολύ σε πρακτικές περιπτώσεις. Για παράδειγμα, όταν $k=3$ και $n=7$, από τον τύπο (4.19), λαμβάνουμε:

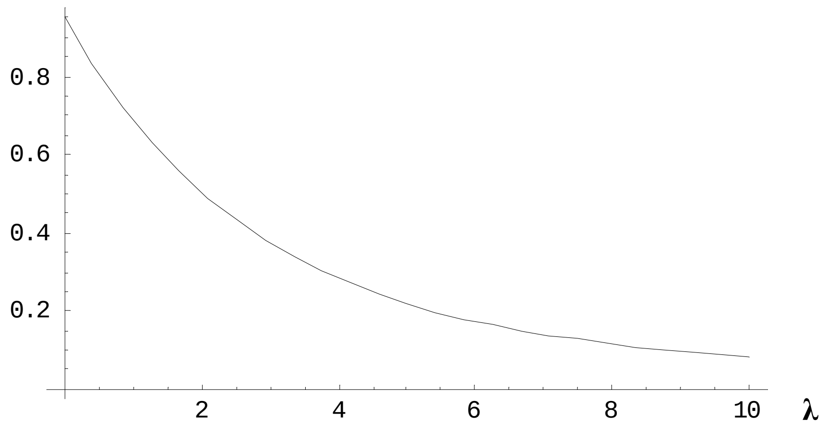
$$R_{\phi_0}(\theta,\sigma) - R_{\phi^+_{JS}}(\theta,\sigma) = \frac{2}{3}\,\frac{D(\sqrt{\lambda/2}) - e^{-\frac{3}{7}\lambda}D(\sqrt{\lambda/14})}{\sqrt{\lambda/2}} + \frac{2e^{-\frac{3}{7}\lambda}}{7\sqrt{7}}\Big(5+\frac{\lambda}{7}\Big)$$

Για $\lambda = 0$, η τιμή του κέρδους είναι:

$$\frac{2}{3} + \frac{16}{21\sqrt{7}} \qquad \Big(\text{καθώς} \quad \lim_{\lambda\to 0}\frac{D(\lambda)}{\lambda} = 1\Big)$$

το οποίο μπορούμε να συγκρίνουμε με την τιμή του κέρδους του εκτιμητή των $James - Stein$, $2/3$, (υπολογιζόμενη με τη βοήθεια της (4.11), βλ.§4.7.1)

Στο παρακάτω γράφημα παριστάνεται το κέρδος του εκτιμητή ϕ^+_{JS}, όταν k=3 και $n = 7$



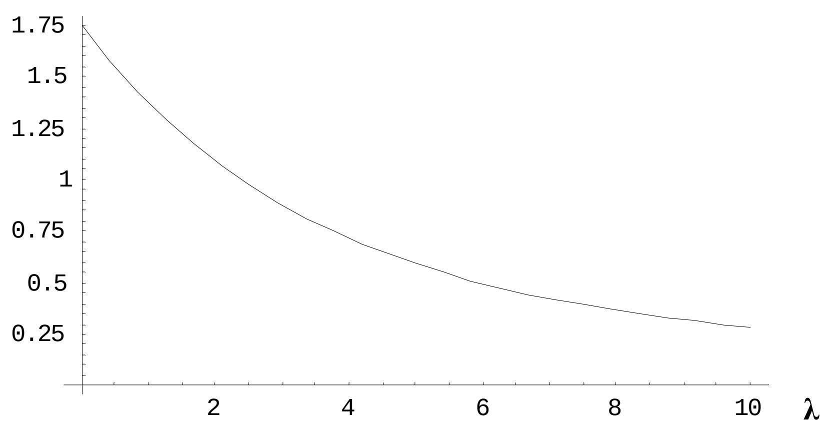
k = 3

Και όταν το $k = 4$ και $n = 8$, χρησιμοποιώντας την (4.18), το κέρδος είναι:

$$R_{\phi_0}(\theta, \sigma) - R_{\phi^+_{JS}}(\theta, \sigma) = \frac{8}{3\lambda}\Big(1 - e^{-3\lambda/8}\Big) + \frac{e^{-3\lambda/8}}{4}\Big(3 + \frac{\lambda}{8}\Big)$$

Όταν το $\lambda = 0$, το κέρδος είναι $7/4$ ενώ εκείνο του εκτιμητή ϕ_{JS} είναι $4/3$.

Στο παρακάτω γράφημα παριστάνεται το κέρδος του εκτιμητή ϕ^+_{JS}, όταν k=4 και $n = 8$



k = 4

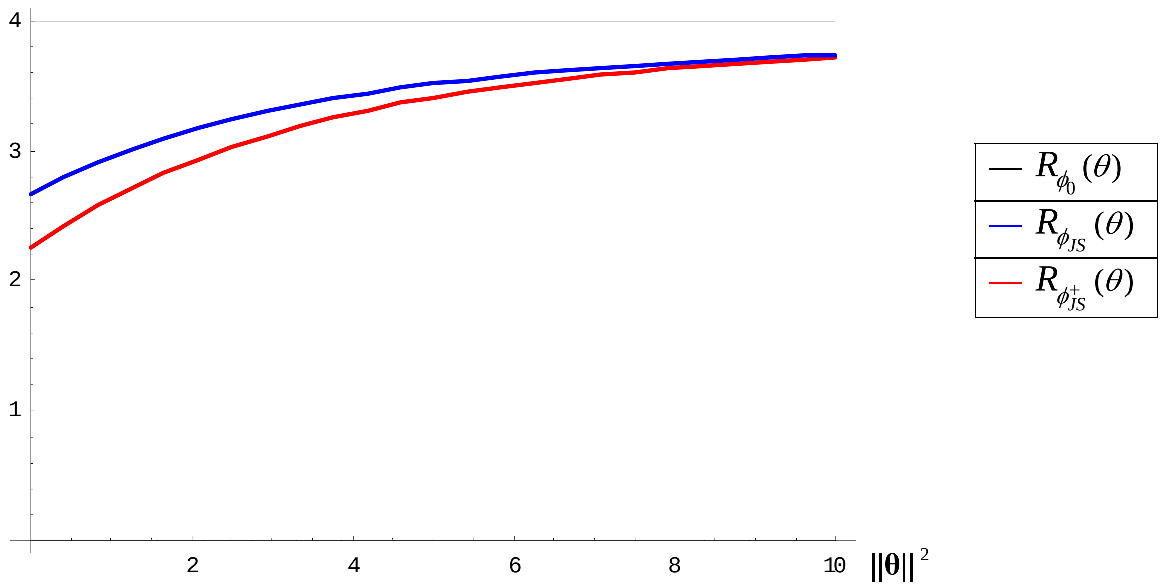
Τέλος δίνουμε ένα ενδεικτικό γράφημα των κινδύνων των εκτιμητών ϕ_0, ϕ_{JS} και ϕ^+_{JS}, συναρτήσει του $\|\theta\|^2$, θεωρώντας το $\sigma = 1$, για $k = 4$ και $n = 8$. Συγκεκριμένα, στην πέρίπτωση αυτή έχουμε:

$$R_{\phi^+_{JS}}(\theta) = 4 - \frac{8}{3\|\theta\|^2}\Big(1 - e^{-3\|\theta\|^2/8}\Big) - \frac{e^{-3\|\theta\|^2/8}}{4}\Big(3 + \frac{\|\theta\|^2}{8}\Big)$$

και

$$R_{\phi_{JS}}(\theta) = 4 - \frac{8}{3\|\theta\|^2}\Big(1 - e^{-\|\theta\|^2/2}\Big)$$

Στο γράφημα αυτό φαίνεται η "βελτίωση" την οποία επιφέρει ο ϕ_{JS}^{+} συγκριτικά με τον ϕ_{JS}.



ΕΦΑΡΜΟΓΗ:

Ένα άνω φράγμα του κέρδους φερόμενο από έναν εκτιμητή με συρρικνωτή

Έχουμε μελετήσει στις προηγούμενες παραγράφους υπό ποιές συνθήκες είναι δυνατή η βελτίωση ενός εκτιμητή με δεδομένο συρρικνωτή. Θα δούμε παρακάτω μία μέθοδο που μας επιτρέπει να εκτιμήσουμε την τιμή ενός εκτιμητή ϕ^1 (με την έννοια ενός δεδομένου τετραγωνικού κινδύνου) εκτιμώντας, με προσομοίωση, το άνω φράγμα του κέρδους G, το οποίο μπορεί να επιφέρει ένας εκτιμητής με συρρικνωτή, ϕ, αναφορικά με τον ε.ε.τ ϕ_0 ($G = R_{\phi_0} - R_{\phi}$). Μπορούμε λοιπόν, να συγκρίνουμε (πάντοτε με προσομοίωση) το κέρδος του εκτιμητή ϕ^1 με το φράγμα αυτό.

Έστω E διανυσματικός χώρος, διαστάσεως n, επί του $\mathbb{R}$ και έστω $y \in E$ ακολουθώντας ένα νόμο με ελλειπτική συμμετρία $P_{\theta,w}$ παραμέτρου θέσεως $\theta \in \Theta$ (Θ γνωστός διανυσματικός υπόχωρος του E, διαστάσεως k, $k \geq 3$) και παραμέτρου διακύμανσης w, διγραμμική συμμετρική μορφή θετικά ορισμένη επί του E^* (δυϊκός του E). Υποθέτουμε ότι η w είναι γνωστή εκτός από έναν πολλαπλασιαστικό παράγοντα, έστω $w = \sigma^2 v$ ($\sigma^2 (> 0)$ άγνωστο, v γνωστή δ.σ.μ, θετικά ορισμένη επί του E^*).

Υποθέτουμε ότι: $(H): \quad \displaystyle\int_E \|y\|^2_{v^{-1}} P_{0,\sigma^2 v}(dy) < +\infty$

Στη συνέχεια διατηρούμε τους συμβολισμούς και τις έννοιες των προηγούμενων κεφαλαίων ενώ η χρησιμοποιούμενη απώλεια είναι η γενική τετραγωνική απώλεια:

$$\sigma^{-2}\overline{q}(\hat{\theta} - \theta)$$

Έστω $\theta \in \Theta$ σταθεροποιημένο. Μεταξύ των εκτιμητών του θ της μορφής:

$$\phi(y) = \Big(1 - h(\phi_0(y), s^2(y))c\Big)\phi_0(y) \tag{1}$$

όπου c είναι ένας ενδομορφισμός του Θ και h μία μετρήσιμη συνάρτηση από τον $\Theta \times \mathbb{R}_+$ στον $\mathbb{R}_+$, υπάρχει ένας εκτιμητής (της μορφής αυτής) ο οποίος ελαχιστοποιεί τον τετραγωνικό κίνδυνο $R_\phi(\theta, \sigma)$ που συνδέεται με τη θετική, διγραμμική, συμμετρική μορφή q στο (θ, σ^2), για οποιοδήποτε σ^2:

$$R_\phi(\theta, \sigma) = \frac{1}{\sigma^2}\mathbb{E}_{\theta,\sigma}[\overline{q}(\phi(y) - \theta)] = \frac{1}{\sigma^2}\int_E \overline{q}(\phi(y) - \theta) P_{\theta,\sigma^2 v}(dy) \ .$$

Ο εκτιμητής αυτός είναι εκείνος που συνδέεται με την:

$$h^*_\theta\Big(\phi_0(y), s^2(y)\Big) = \frac{q\Big(c(\phi_0(y)), \phi_0(y) - \theta\Big)}{\overline{q}\Big(c(\phi_0(y))\Big)}$$

Πράγματι, για τον κίνδυνο που συνδέεται με την q έχουμε:

$$\begin{aligned}
&R_{\phi}(\theta,\sigma) - R_{\phi_0}(\theta,\sigma)\\
&= \frac{1}{\sigma^2}\mathbb{E}_{\theta,\sigma}\Big[\overline{q}\Big(\phi_0(y) - \theta - h(\phi_0(y), s^2(y))c(\phi_0(y))\Big)\Big] - \frac{1}{\sigma^2}\mathbb{E}_{\theta,\sigma}\Big[\overline{q}\Big(\phi_0(y) - \theta\Big)\Big]\\
&= \frac{1}{\sigma^2}\mathbb{E}_{\theta,\sigma}\Big[h^2\Big(\phi_0(y), s^2(y)\Big)\overline{q}\Big(c(\phi_0(y))\Big)\Big] - \frac{2}{\sigma^2}\mathbb{E}_{\theta,\sigma}\Big[h\Big(\phi_0(y), s^2(y)\Big)q\Big(c(\phi_0(y)), \phi_0(y) - \theta\Big)\Big]
\end{aligned}$$

Για κάθε $(\phi_0(y), s^2(y))$, η h^*_θ προφανώς μεγιστοποιεί την έκφραση:

$$h\Big(\phi_0(y), s^2(y)\Big)\Big\{2q\Big(c(\phi_0(y)), \phi_0(y) - \theta\Big) - h\Big(\phi_0(y), s^2(y)\Big)\overline{q}\Big(c(\phi_0(y))\Big)\Big\}$$

Ο κίνδυνος του εκτιμητή ϕ^*_θ που συνδέεται με την h^*_θ στο (θ,σ), γράφεται λοιπόν:

$$\begin{aligned}
R_{\phi^*_\theta}(\theta,\sigma) &= R_{\phi_0}(\theta,\sigma) + \frac{1}{\sigma^2}\mathbb{E}_{\theta,\sigma}\left[\frac{q^2\Big(c(\phi_0(y)), \phi_0(y) - \theta\Big)\overline{q}\Big(c(\phi_0(y))\Big)}{\overline{q}^2\Big(c(\phi_0(y))\Big)}\right]\\
&\quad - \frac{2}{\sigma^2}\mathbb{E}_{\theta,\sigma}\left[\frac{q^2\Big(c(\phi_0(y)), \phi_0(y) - \theta\Big)}{\overline{q}\Big(c(\phi_0(y))\Big)}\right]\\
&= R_{\phi_0}(\theta,\sigma) - \frac{1}{\sigma^2}\mathbb{E}_{\theta,\sigma}\left[\frac{q^2\Big(c(\phi_0(y)), \phi_0(y) - \theta\Big)}{\overline{q}\Big(c(\phi_0(y))\Big)}\right]
\end{aligned}$$

Από την ανισότητα του *Schwarz*, έχουμε:

$$\begin{aligned}
\mathbb{E}_{\theta,\sigma}\left[\frac{q^2\Big(c(\phi_0(y)), \phi_0(y) - \theta\Big)}{\overline{q}\Big(c(\phi_0(y))\Big)}\right] &\leq \mathbb{E}_{\theta,\sigma}\left[\frac{\overline{q}\Big(c(\phi_0(y))\Big)\overline{q}\Big(\phi_0(y) - \theta\Big)}{\overline{q}\Big(c(\phi_0(y))\Big)}\right]\\
&= \mathbb{E}_{\theta,\sigma}\Big[\overline{q}\Big(\phi_0(y) - \theta\Big)\Big]\\
&= \int_E \overline{q}\Big(\phi_0(y) - \theta\Big)P_{\theta,\sigma^2 v}(dy)\\
&= \int_E \overline{q}\Big(\phi_0(y - \theta)\Big)P_{\theta,\sigma^2 v}(dy)\\
&= \int_E \overline{q}\Big(\phi_0(x)\Big)P_{0,\sigma^2 v}(dx)\\
&\leq \max_{1\leq i\leq k} q_i \int_E \|\phi_0(x)\|^2_{v^{-1}}P_{0,\sigma^2 v}(dx)\\
&\leq \max_{1\leq i\leq k} q_i \int_E \|x\|^2_{v^{-1}}P_{0,\sigma^2 v}(dx)
\end{aligned}$$

το οποίο είναι πεπερασμένο λόγω της (H). Άρα, ο κίνδυνος στο (θ,σ) του ϕ^*_θ είναι πεπερασμένος.

Όμως, όπως μπορούμε να δούμε στο 1^o γράφημα παρακάτω, ο εκτιμητής $\phi^*_{\theta_0}$ που συνδέεται με την $h^*_{\theta_0}$, για σταθεροποιημένο θ_0, δεν κυριαρχεί τον ε.ε.τ ϕ_0 ομοιομόρφως ως προς (θ,σ).

Πράγματι, περιοριζόμαστε στην περίπτωση, όπου $c = id_\Theta$, $q = v_\Theta^{-1}$ (και $\sigma = 1$) και θεωρούμε τον εκτιμητή:

$$\phi^*_{\theta_0}(y) = \Big(1 - h^*_{\theta_0}(\phi_0(y), s^2(y))\Big)\phi_0(y) \quad , \quad \text{για} \quad \theta_0\Big(\in \Theta\Big) \ \text{σταθερό}$$

όπου

$$\begin{aligned} h^*_{\theta_0}\Big(\phi_0(y), s^2(y)\Big) &= \frac{v_\Theta^{-1}\Big(\phi_0(y), \phi_0(y) - \theta_0\Big)}{\overline{v_\Theta^{-1}}\Big(\phi_0(y)\Big)} \\ &= \frac{\overline{v_\Theta^{-1}}\Big(\phi_0(y)\Big) - v_\Theta^{-1}\Big(\phi_0(y), \theta_0\Big)}{\overline{v_\Theta^{-1}}\Big(\phi_0(y)\Big)} \\ &= 1 - \frac{v_\Theta^{-1}\Big(\phi_0(y), \theta_0\Big)}{\overline{v_\Theta^{-1}}\Big(\phi_0(y)\Big)} \end{aligned}$$

ή

$$1 - h^*_{\theta_0}\Big(\phi_0(y), s^2(y)\Big) = \frac{v_\Theta^{-1}\Big(\phi_0(y), \theta_0\Big)}{\overline{v_\Theta^{-1}}\Big(\phi_0(y)\Big)}$$

Τότε ο κίνδυνος του παραπάνω εκτιμητή, σ' ένα τυχαίο $\theta \in \Theta$, είναι:

$$\begin{aligned} R_{\phi^*_{\theta_0}}(\theta) &= \mathbb{E}_\theta\Big[\overline{q}(\phi^*_{\theta_0}(y) - \theta)\Big] \\ &= \mathbb{E}_\theta\Big[\overline{q}\Big(\frac{v_\Theta^{-1}(\phi_0(y), \theta_0)}{\overline{v_\Theta^{-1}}(\phi_0(y))}\ \phi_0(y) - \theta\Big)\Big] \\ &= \overline{q}(\theta) + \mathbb{E}_\theta\Big[\Big(\frac{v_\Theta^{-1}(\phi_0(y), \theta_0)}{\overline{v_\Theta^{-1}}(\phi_0(y))}\Big)^2 \overline{q}(\phi_0(y))\Big] - 2\mathbb{E}_\theta\Big[\frac{v_\Theta^{-1}(\phi_0(y), \theta_0)}{\overline{v_\Theta^{-1}}(\phi_0(y))}\ q(\phi_0(y), \theta)\Big] \\ (q = v_\Theta^{-1}) \quad &= \overline{v_\Theta^{-1}}(\theta) + \mathbb{E}_\theta\Bigg[\frac{\Big(v_\Theta^{-1}(\phi_0(y), \theta_0)\Big)^2}{\overline{v_\Theta^{-1}}(\phi_0(y))}\Bigg] - 2\mathbb{E}_\theta\Bigg[\frac{v_\Theta^{-1}(\phi_0(y), \theta_0) v_\Theta^{-1}(\phi_0(y), \theta)}{\overline{v_\Theta^{-1}}(\phi_0(y))}\Bigg] \end{aligned}$$

Απ' όπου, για $\theta = \theta_0$, λαμβάνουμε ότι:

$$R_{\phi^*_{\theta_0}}(\theta_0) = \overline{v_\Theta^{-1}}(\theta_0) - \mathbb{E}_{\theta_0}\Bigg[\frac{\Big(v_\Theta^{-1}(\phi_0(y), \theta_0)\Big)^2}{\overline{v_\Theta^{-1}}(\phi_0(y))}\Bigg]$$

Παρατηρούμε ότι οι στατιστικές συναρτήσεις:

$$T_1(y) = \frac{\Big(v_\Theta^{-1}(\phi_0(y), \theta_0)\Big)^2}{\overline{v_\Theta^{-1}}(\phi_0(y))} \qquad \text{και} \qquad T_2(y) = \frac{v_\Theta^{-1}(\phi_0(y), \theta_0)\ v_\Theta^{-1}(\phi_0(y), \theta)}{\overline{v_\Theta^{-1}}(\phi_0(y))}$$

είναι ελεύθερες υπό ομοθεσία $h_r(x) = rx$ $(r > 0)$.
(δηλ. $\forall x \in E$, $T_i \circ h_r(x) = T_i(x)$, $i = 1, 2$)

Λοιπόν, αν $\mathcal{U}_{v,r,\theta}$ είναι ο ομοιόμορφος νόμος επί της σφαίρας $S_{v,r,\theta} = \{x \in E : \|x-\theta\|^2 = r^2\}$, κέντρου θ και ακτίνας r, έχουμε:

$$\begin{aligned} \Big(\mathcal{U}_{v,r,\theta}\Big)_{T_i}(B) &= \mathcal{U}_{v,r,\theta}\Big(T_i^{-1}(B)\Big) = \Big(\Big(\mathcal{U}_{v,1,\theta}\Big)_{h_r}\Big)\Big(T_i^{-1}(B)\Big) \\ &= \Big(\mathcal{U}_{v,1,\theta}\Big)_{T_i \circ h_r}(B) = \Big(\mathcal{U}_{v,1,\theta}\Big)_{T_i}(B) \end{aligned}$$

Επομένως, λαμβάνοντας υπόψη ότι ένας v-ελλειπτικά συμμετρικός νόμος γύρω από το θ είναι ένα μίγμα ομοιομόρφων νόμων (βλ.[22]), έχουμε:

$$\begin{aligned} \Big(\mathcal{U}_{v,1,\theta}\Big)_{T_i}(B) &= \int_{\mathbb{R}_+} \Big(\mathcal{U}_{v,r,\theta}\Big)_{T_i}(B) P_{\|\cdot\|}(dr) \\ &= \int_{\mathbb{R}_+} \Big(\int_{S_{v,r,\theta}} \mathbf{1}_{T_i^{-1}(B)} d\mathcal{U}_{v,r,\theta}\Big) P_{\|\cdot\|}(dr) \\ &= P_{\theta,v}\Big(T_i^{-1}(B)\Big) = \Big(P_{\theta,v}\Big)_{T_i}(B) \end{aligned}$$

δηλαδή, ο νόμος της T_i (εικόνα του $P_{\theta,v}$ υπό την T_i, για οποιονδήποτε $P_{\theta,v}$) είναι η εικόνα του ομοιόμορφου νόμου $\mathcal{U}_{v,1,\theta}$ μέσω της T_i. Χωρίς, λοιπόν, περιορισμό της γενικότητας λαμβάνουμε:

$$P_{\theta,v} = N_E(\theta, v)$$

Αναλυτικός τύπος των κινδύνων $R_{\phi^*_{\theta_0}}(\theta_0)$ ***και*** $R_{\phi^*_{\theta_0}}(\theta)$.

Με τη βοήθεια του Θεωρήματος 1(βλ.Παράρτημα), έχουμε ότι:

$$\begin{aligned} &\mathbb{E}_\theta\Big[\frac{\Big(v_\Theta^{-1}(\phi_0(y),\theta_0)\Big)^2}{\overline{v_\Theta^{-1}}(\phi_0(y))}\Big] = \int_E \frac{\Big(v_\Theta^{-1}(\phi_0(y),\theta_0)\Big)^2}{\overline{v_\Theta^{-1}}(\phi_0(y))} N_E(\theta, v; dy) \\ &= \int_{\mathbb{R}^{n-k}} \Big(\int_{\mathbb{R}^k} \frac{{}^t\theta_0 x {}^t x \theta_0}{{}^t x x}\ N_k(\theta, I_k; dx)\Big) N_{n-k}(0, I_{n-k}; dz) \\ &= {}^t\theta_0 \Big(\int_{\mathbb{R}^k} \frac{x {}^t x}{{}^t x x}\ N_k(\theta, I_k; dx)\Big)\theta_0 \\ &= {}^t\theta_0\Big(\int_{\mathbb{R}_+} \frac{1}{u}\ \mathcal{X}^2_{k+2}(2\lambda; du)\Big) I_k\theta_0 + {}^t\theta_0\Big(\int_{\mathbb{R}_+} \frac{1}{u}\ \mathcal{X}^2_{k+4}(2\lambda; du)\Big)\theta {}^t\theta\theta_0 \\ &= \|\theta_0\|^2\Big(\int_{\mathbb{R}_+} \frac{1}{u}\ \mathcal{X}^2_{k+2}(2\lambda; du)\Big) + {}^t\theta_0\theta {}^t\theta\theta_0\Big(\int_{\mathbb{R}_+} \frac{1}{u}\ \mathcal{X}^2_{k+4}(2\lambda; du)\Big) \\ &= \|\theta_0\|^2\Big(2^{-1}\int_{\mathbb{N}_0} \frac{\Gamma(\frac{k+2}{2}-1+m)}{\Gamma(\frac{k+2}{2}+m)}\ \pi(\lambda; dm)\Big) + {}^t\theta_0\theta {}^t\theta\theta_0\Big(2^{-1}\int_{\mathbb{N}_0} \frac{\Gamma(\frac{k+4}{2}-1+m)}{\Gamma(\frac{k+4}{2}+m)}\ \pi(\lambda; dm)\Big) \\ &= \frac{\|\theta_0\|^2}{2}\Big(\int_{\mathbb{N}_0} \frac{\Gamma(\frac{k}{2}+m)}{(\frac{k}{2}+m)\Gamma(\frac{k}{2}+m)}\ \pi(\lambda; dm)\Big) + {}^t\theta_0\theta {}^t\theta\theta_0\Big(\frac{1}{2}\int_{\mathbb{N}_0} \frac{\Gamma(\frac{k}{2}+m+1)}{(\frac{k}{2}+m+1)\Gamma(\frac{k}{2}+m+1)}\ \pi(\lambda; dm)\Big) \\ &= \frac{\|\theta_0\|^2}{2}\Big(\int_{\mathbb{N}_0} \frac{1}{\frac{k}{2}+m}\ \pi(\lambda; dm)\Big) + {}^t\theta_0\theta {}^t\theta\theta_0\Big(\frac{1}{2}\int_{\mathbb{N}_0} \frac{1}{\frac{k}{2}+m+1}\ \pi(\lambda; dm)\Big) \end{aligned}$$

Επειδή, όπως είδαμε και στην §4.7.1,

$$\int_{\mathbb{N}_0} \frac{1}{\frac{k}{2}+m}\ \pi(\lambda; dm) = \frac{e^{-\lambda}}{\lambda^{k/2}} \int_0^\lambda x^{\frac{k}{2}-1} e^x dx$$

και

$$\int_{\mathbb{N}_0} \frac{1}{\frac{k}{2}+m+1}\ \pi(\lambda; dm) = \frac{e^{-\lambda}}{\lambda^{\frac{k}{2}+1}} \int_0^\lambda x^{\frac{k}{2}} e^x dx\ ,$$

προκύπτει ότι:

$$\mathbb{E}_\theta\Bigg[\frac{\Big(v_\Theta^{-1}(\phi_0(y),\theta_0)\Big)^2}{\overline{v_\Theta^{-1}}(\phi_0(y))}\Bigg] = \frac{\|\theta\|^2 e^{-\lambda}}{2\lambda^{k/2}} \int_0^\lambda x^{\frac{k}{2}-1} e^x dx + {}^t\theta_0\theta^t\theta\theta_0\Bigg(\frac{e^{-\lambda}}{2\lambda^{\frac{k}{2}+1}} \int_0^\lambda x^{\frac{k}{2}} e^x dx\Bigg)$$

Επίσης,

$$\begin{aligned}
&\mathbb{E}_\theta\Bigg[\frac{v_\Theta^{-1}(\phi_0(y),\theta_0)v_\Theta^{-1}(\phi_0(y),\theta)}{\overline{v_\Theta^{-1}}(\phi_0(y))}\Bigg] = \int_{\mathbb{R}^k} \frac{{}^t\theta_0 x\, {}^t x\theta}{{}^t x x}\ N_k(\theta, I_k; dx)\\
&= {}^t\theta_0\Bigg(\int_{\mathbb{R}^k} \frac{x\,{}^t x}{{}^t x x}\ N_k(\theta, I_k; dx)\Bigg)\theta\\
&= {}^t\theta_0\Bigg(\int_{\mathbb{R}_+} \frac{1}{u}\ \mathcal{X}^2_{k+2}(2\lambda; du)\Bigg) I_k\theta + {}^t\theta_0\Bigg(\int_{\mathbb{R}_+} \frac{1}{u}\ \mathcal{X}^2_{k+4}(2\lambda; du)\Bigg)\theta^t\theta\theta\\
&= {}^t\theta_0\theta\Bigg(\frac{e^{-\lambda}}{2\lambda^{k/2}} \int_0^\lambda x^{\frac{k}{2}-1} e^x dx\Bigg) + {}^t\theta_0\theta\Bigg(\frac{\|\theta\|^2 e^{-\lambda}}{2\lambda^{\frac{k}{2}+1}} \int_0^\lambda x^{k/2} e^x dx\Bigg)
\end{aligned}$$

Συνεπώς,

$$\begin{aligned}
R_{\phi^*_{\theta_0}}(\theta) \ &= \ \|\theta\|^2 + \|\theta_0\|^2 \frac{e^{-\lambda}}{2\lambda^{k/2}} \int_0^\lambda x^{\frac{k}{2}-1} e^x dx + {}^t\theta_0\theta^t\theta\,\theta_0\Bigg(\frac{e^{-\lambda}}{2\lambda^{\frac{k}{2}+1}} \int_0^\lambda x^{\frac{k}{2}} e^x dx\Bigg)\\
&\quad - 2\,{}^t\theta_0\theta\Bigg(\frac{e^{-\lambda}}{2\lambda^{k/2}} \int_0^\lambda x^{\frac{k}{2}-1} e^x dx\Bigg) - 2\,{}^t\theta_0\theta\|\theta\|^2\Bigg(\frac{e^{-\lambda}}{2\lambda^{\frac{k}{2}+1}} \int_0^\lambda x^{\frac{k}{2}} e^x dx\Bigg)
\end{aligned}$$

δηλαδή,

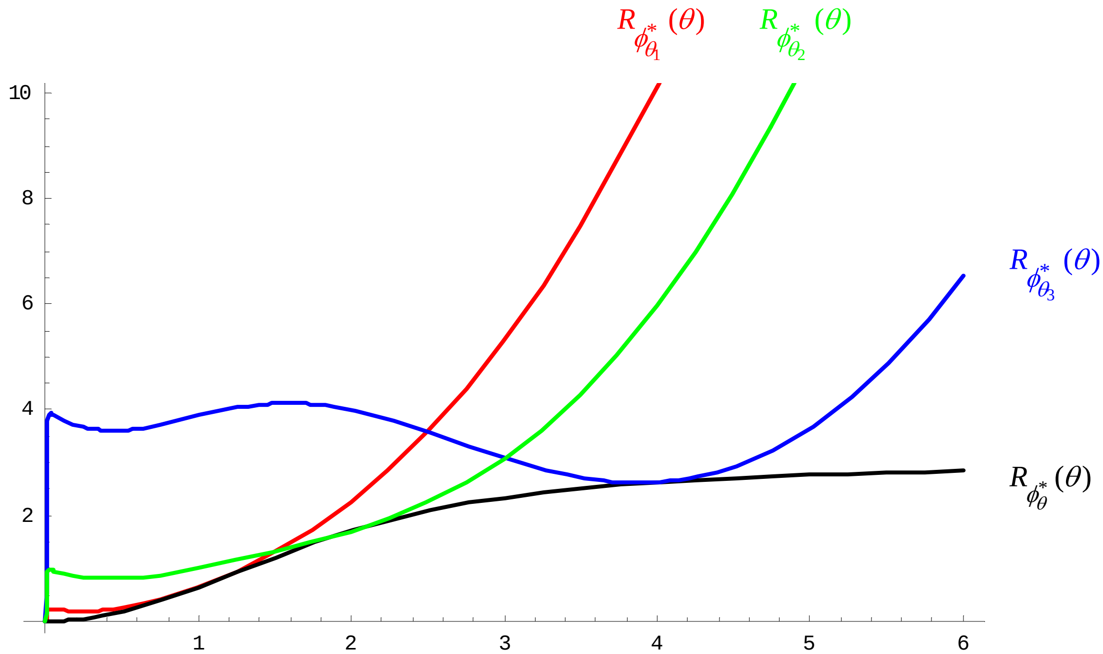
$$\begin{aligned}
R_{\phi^*_{\theta_0}}(\theta) \ &= \ \|\theta\|^2 - \Big(2\,{}^t\theta_0\theta - \|\theta_0\|^2\Big)\frac{e^{-\lambda}}{2\lambda^{k/2}} \int_0^\lambda x^{\frac{k}{2}-1} e^x dx\\
&\quad - {}^t\theta_0\theta\Big(2\|\theta\|^2 - {}^t\theta_0\theta\Big)\frac{e^{-\lambda}}{2\lambda^{\frac{k}{2}+1}} \int_0^\lambda x^{k/2} e^x dx \qquad (2)
\end{aligned}$$

και

$$R_{\phi^*_{\theta_0}}(\theta_0) = \|\theta_0\|^2 - \|\theta_0\|^2 \frac{e^{-\lambda}}{2\lambda^{k/2}} \int_0^\lambda x^{\frac{k}{2}-1} e^x dx - \|\theta_0\|^4 \frac{e^{-\lambda}}{2\lambda^{\frac{k}{2}+1}} \int_0^\lambda x^{k/2} e^x dx \qquad (3)$$

Τα αποτελέσματα αυτά μας δίνουν τη δυνατότητα να δώσουμε το παρακάτω γράφημα. Στο γράφημα αυτό φαίνονται οι γραφικές παραστάσεις των $R_{\phi^*_\theta}(\theta)$, $R_{\phi^*_{\theta_1}}(\theta)$, $R_{\phi^*_{\theta_2}}(\theta)$, $R_{\phi^*_{\theta_3}}(\theta)$ συναρτήσει του $\|\theta\|$, για $k=4$, όπου $R_{\phi^*_\theta}(\theta)$ δίνεται από τον τύπο (3) και τα $R_{\phi^*_{\theta_1}}(\theta)$, $R_{\phi^*_{\theta_2}}(\theta)$,

$R_{\phi^*_{\theta_3}}(\theta)$ από τον τύπο (2), αν θεωρήσουμε ότι το $\theta = C\theta_i$ $(i = 1,2,3)$ με $C > 0$ και $\|\theta_1\| = 1$, $\|\theta_2\| = 2$ και $\|\theta_3\| = 4$ αντιστοίχως.



Γράφημα 1^o: Κίνδυνοι των εκτιμητών $\phi^*_{\theta_i}$, $i = 1,2,3$, συναρτήσει του $\|\theta\|$, για $k = 4$.

Παρατηρήσεις:

1) Οι εκτιμητές της μορφής:

$$\phi^*_{\theta_i}(y) = \frac{v_\Theta^{-1}(\phi_0(y), \theta_i)}{\overline{v_\Theta^{-1}}(\phi_0(y))}\, \phi_0(y)$$

για θ_i σταθεροποιημένο $(i = 1,2,3)$ δεν παρουσιάζουν ιδιαίτερο ενδιαφέρον αφού, όπως μπορούμε να δούμε και στο παραπάνω γράφημα, δεν κυριαρχούν ομοιομόρφως, με την έννοια του κινδύνου, τον ε.ε.τ ϕ_0, ο οποίος έχει σταθερό κίνδυνο και ίσο με 4.

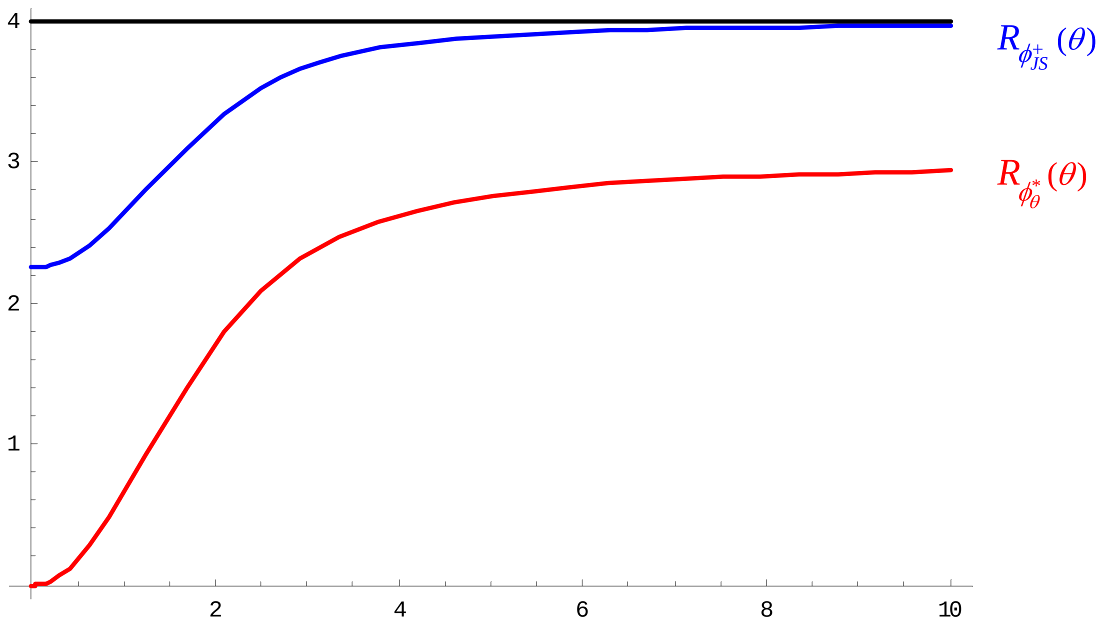
2) Επειδή η καμπύλη του κινδύνου $R_{\phi^*_{\theta_i}}(\theta)$ εφάπτεται εκείνης του $R_{\phi^*_\theta}(\theta)$ (ελάχιστος κίνδυνος) ο εκτιμητής $\phi^*_{\theta_i}$ είναι αποδοτικός αν το θ, αληθινή τιμή της παραμέτρου, είναι πλησίον του θ_i.

Μπορούμε επίσης να θεωρήσουμε την ακολουθία των εκτιμητών

$$\phi^i(y) = \frac{v_\Theta^{-1}(\phi_0(y), \theta^{i-1})}{\overline{v_\Theta^{-1}}(\phi_0(y))}\phi_0(y)$$

όπου $\theta^{i-1} = \phi^{i-1}(y)$ για $i \geq 2$ και θ^0 να είναι μία εκτίμηση του θ, από έναν εκτιμητή με συρρικνωτή (για παράδειγμα εκείνον των $James - Stein$ "θετικού μέρους"). Όμως, οι πραγματοποιούμενες προσομοιώσεις (βλ.[41]) δεν έχουν αποδείξει τη σύγκλιση αυτής της ακολουθίας, οι οποίες από τις πρώτες επαναλήψεις παρουσιάζουν έναν κίνδυνο μεγαλύτερο από εκείνον των $James - Stein$.

Στο παρακάτω γράφημα φαίνεται ο κίνδυνος του εκτιμητή $James - Stein$ "θετικού μέρους" (βλ.§4.7.2) και ο ελάχιστος κίνδυνος $R_{\phi_\theta^*}(\theta)$ ενός εκτιμητή με συρρικνωτή συναρτήσει του $\|\theta\|$, όταν η διάσταση k του Θ είναι 4.



Γράφημα 2^o: Κίνδυνος του εκτιμητή των $James - Stein$ "θετικού μέρους" και ελάχιστος κίνδυνος ενός εκτιμητή με συρρικνωτή ϕ_θ^*, συναρτήσει του $\|\theta\|$, για $k = 4$.

ΠΑΡΑΡΤΗΜΑ

Η ΜΗ-ΚΕΝΤΡΙΚΗ $\mathcal{X}^2$-ΚΑΤΑΝΟΜΗ

Λήμμα 1

Η μη-κεντρική $\mathcal{X}^2$-πιθανότητα με n βαθμούς ελευθερίας και παράμετρο δ, $\mathcal{X}_n^2(\delta)$, είναι η σύνθεση της πιθανότητας Poisson παραμέτρου $\delta/2$, $\pi(\delta/2)$, και της μετάβασης πιθανότητας από τον $\mathbb{N}_0$ στον $\mathbb{R}_+$, η οποία με κάθε k συνδέει την κεντρική $\mathcal{X}^2$-πιθανότητα με $n+2k$ βαθμούς ελευθερίας, $\mathcal{X}_{n+2k}^2$.

Άλλες εκφράσεις του Λήμματος

(1) Αν $X_1, ... X_n$ ανεξάρτητες τ.μ.: $X_i \sim N(\mu_i, 1)$ $i = 1, ..., n$, τότε

$$\sum_{i=1}^{n} X_i^2 \sim \mathcal{X}_n^2(\delta) \quad \text{όπου} \;\; \delta = \sum_{i=1}^{n} \mu_i^2$$

(2) Αν μία τ.μ Y ακολουθεί, δεσμευτικά με μία τ.μ K, το νόμο $\mathcal{X}_{n+2K}^2$ και αν K ακολουθεί την *Poisson* $\pi(\delta/2)$ τότε η Y ακολουθεί την $\mathcal{X}_n^2(\delta)$.

(3) Για κάθε απεικόνιση, f από τον $\mathbb{R}_+$ μέσα στον $\mathbb{R}$, $\mathcal{X}_n^2(\delta)$-ολοκληρώσιμη, έχουμε:

$$\int_{\mathbb{R}_+} f(x)\mathcal{X}_n^2(\delta; dx) = \int_{\mathbb{N}_0} \left(\int_{\mathbb{R}_+} f(x)\mathcal{X}_{n+2k}^2(dx) \right) \pi(\delta/2; dk)$$

Απόδειξη

Η πυκνότητα πιθανότητας η οποία προκύπτει από την σύνθεση της $\pi(\delta/2)$ και της μετάβασης $k \rightsquigarrow \mathcal{X}_{n+2k}^2$ είναι

$$\sum_{k=0}^{+\infty} \frac{e^{-\delta/2}}{k!} \left(\frac{\delta}{2}\right)^k \frac{y^{\frac{n}{2}+k-1}}{\Gamma(\frac{n}{2}+k)} \, \frac{e^{-y/2}}{2^{\frac{n}{2}+k}}$$

η οποία δεν είναι άλλη από την πυκνότητα της $\mathcal{X}_n^2(\delta)$ (βλ. [21]).

□

Πόρισμα 1

Έστω $u \in \mathbb{R}$, η ροπή τάξεως u της μη-κεντρικής $\mathcal{X}_n^2$-πιθανότητας, $\mathcal{X}_n^2(\delta)$, είναι πεπερασμένη αν και μόνον αν $u > -n/2$. Σ'αυτή την περίπτωση,

$$\int_{\mathbb{R}_+} x^u \mathcal{X}_n^2(\delta; dx) = 2^u \int_{\mathbb{N}_0} \frac{\Gamma(\frac{n}{2}+u+k)}{\Gamma(\frac{n}{2}+k)} \, \pi(\delta/2; dk)$$

Απόδειξη

Από την έκφραση (3) του Λήμματος 1, έχουμε:

$$\int_{\mathbb{R}_+} x^u \mathcal{X}_n^2(\delta;dx) = \int_{\mathbb{N}_0}\Big(\int_{\mathbb{R}_+} x^u \mathcal{X}_{n+2k}^2(dx)\Big)\pi(\frac{\delta}{2};dk) = \sum_{k=0}^{+\infty}\Big(\int_{\mathbb{R}_+} x^u \mathcal{X}_{n+2k}^2(dx)\Big)\frac{e^{-\frac{\delta}{2}}(\frac{\delta}{2})^k}{k!} \tag{1}$$

όπου $\forall k \in N_0, \quad \int_{\mathbb{R}_+} x^u \mathcal{X}_{n+2k}^2(dx) \ \in \ (0,+\infty]$

• αν $u \leq -\frac{n}{2}$, τότε

$$\int_{\mathbb{R}_+} x^u \mathcal{X}_n^2(dx) = \int_{\mathbb{R}_+} x^u \ \frac{1}{\Gamma(\frac{n}{2})2^{\frac{n}{2}}} \ x^{\frac{n}{2}-1}e^{-\frac{x}{2}}dx = \frac{1}{\Gamma(\frac{n}{2})2^{\frac{n}{2}}}\int_{\mathbb{R}_+} x^{(u+\frac{n}{2})-1}e^{-\frac{x}{2}}dx$$

Θέτοντας $a = u + \frac{n}{2}$, το τελευταίο ολοκλήρωμα γράφεται:

$$\int_{\mathbb{R}_+} x^a e^{-\frac{x}{2}}dx = \int_0^1 x^a e^{-\frac{x}{2}}dx + \int_1^{+\infty} x^a e^{-\frac{x}{2}}dx = I_1 + I_2 \quad \text{όπου } a \leq -1$$

Επειδή $e^{-\frac{x}{2}} \downarrow$ στο $[0,1]$,

$$I_1 \geq \int_0^1 x^a e^{-\frac{1}{2}}dx = +\infty$$

ενώ

$$I_2 < +\infty$$

Επομένως, $\int_{\mathbb{R}_+} x^u \mathcal{X}_n^2(dx) = +\infty$ δηλαδή ο μηδενικός όρος στην (1) απειρίζεται, άρα

$$\int_{\mathbb{R}_+} x^u \mathcal{X}_n^2(\delta;dx) = +\infty$$

• αν $u > -\frac{n}{2}$, $\forall k \in \mathbb{N}_0$,

$$\begin{aligned}
\int_{\mathbb{R}_+} x^u \mathcal{X}_{n+2k}^2(dx) &= \int_{\mathbb{R}_+} x^u \ \frac{1}{\Gamma(k+\frac{n}{2})2^{k+\frac{n}{2}}} \ x^{k+\frac{n}{2}-1}e^{-\frac{x}{2}}dx \\
&= \frac{1}{\Gamma(k+\frac{n}{2})2^{k+\frac{n}{2}}}\int_{\mathbb{R}_+} x^{(u+k+\frac{n}{2})-1}e^{-\frac{x}{2}}dx \\
&= \frac{1}{\Gamma(k+\frac{n}{2})2^{k+\frac{n}{2}}}\cdot\Gamma(u+k+\frac{n}{2})2^{u+k+\frac{n}{2}} \\
&= 2^u \ \frac{\Gamma(u+k+\frac{n}{2})}{\Gamma(\frac{n}{2}+k)} \qquad (2)
\end{aligned}$$

και η ζητούμενη σχέση προκύπτει άμεσα από τις (1) και (2).

Μένει να επαληθεύσουμε σε αυτήν την περίπτωση ότι $\int_{\mathbb{R}_+} x^u \mathcal{X}_n^2(\delta;dx) < +\infty$. Για να το αποδείξουμε χρησιμοποιούμε το γεγονός ότι:

$$\forall t \geq 0, \quad \mathcal{X}_n^2(\ \delta,\ [0,t]\ ) \leq \mathcal{X}_n^2(\ 0,\ [0,t]\ )$$

Πράγματι, έστω $M = \ {}^t(X_1,...,X_n)$ διανυσματική τυχαία μεταβλητή με $X_i \sim N(\mu_i,1)$, $i = 1,...,n$, ανεξάρτητες. Αν $Z = \sum_{i=1}^n X_i^2$ τότε $Z \sim \mathcal{X}_n^2(\delta)$ όπου $\delta = \sum_{i=1}^n \mu_i^2$. Μέσω ενός ορθογώνιου μετασχηματισμού η Z γράφεται

$$Z = U + V \quad \text{όπου} \quad U = \sum_{i=1}^{n-1} {Y_i}^2 \text{ και } V = {Y_n}^2$$

με $Y_i \sim N(0,1),\ i=1,...,n-1$ και $Y_n \sim N(\sqrt{\delta},1)$. Για κάθε $z \geq 0$, έχουμε:

$$\begin{aligned} F_V(z) &= P(V \leq z) = P({Y_n}^2 \leq z) = P(-\sqrt{z} \leq Y_n \leq \sqrt{z}) \\ &= P(-\sqrt{z}-\sqrt{\delta} \leq Y_n - \sqrt{\delta} \leq \sqrt{z}-\sqrt{\delta}) \\ &= \Phi(\sqrt{z}-\sqrt{\delta}) - \Phi(-\sqrt{z}-\sqrt{\delta}) \end{aligned}$$

όπου Φ η συνάρτηση κατανομής της $N(0,1)$.

Παρατηρούμε ότι, καθώς $\delta \uparrow,\ F_V(z) \downarrow$. Άρα, για κάθε $z \geq 0$, έχουμε:

$$\begin{aligned} F_Z(z) = P(U+V \leq z) &= \int_0^z \Big(\int_0^{-u+z} f_{(U,V)}(u,v)dv \Big) du \\ &= \int_0^z \Big(\int_0^{-u+z} f_V(v)dv \Big) f_U(u)du \\ &= \int_0^z F_V(z-u) f_U(u)du \end{aligned}$$

το οποίο φθίνει αν το δ αυξάνει. Συνεπώς, επειδή $Z \sim \mathcal{X}_n^2(\delta)$,

$$\forall z \geq 0, \quad \mathcal{X}_n^2(\delta,[0,z]) \leq \mathcal{X}_n^2(0,[0,z])$$

Επειδή $x^u \geq 0$, προκύπτει:

$$\int_0^{+\infty} x^u \mathcal{X}_n^2(\delta;dx) \leq \int_0^{+\infty} x^u \mathcal{X}_n^2(0;dx) < +\infty$$

και το πόρισμα αποδείχτηκε.

$\square$

Λήμμα 2

Αν $Z\sim \mathcal{X}_n^2(\delta)$ τότε η χαρακτηριστική της συνάρτηση δίνεται από τον τύπο:

$$\phi_Z(t) = (1-2it)^{-n/2} e^{it\delta/(1-2it)}$$

(βλ.[29], σελ.24)

Από το Λήμμα 2, άμεσα προκύπτει ότι:

Λήμμα 3

Αν Z_1, Z_2 ανεξάρτητες τ.μ. με $Z_1 \sim \mathcal{X}_{n_1}^2(\delta_1)$ και $Z_2 \sim \mathcal{X}_{n_2}^2(\delta_2)$ τότε $Z_1+Z_2 \sim \mathcal{X}_{n_1+n_2}^2(\delta_1+\delta_2)$

Λήμμα 4

Αν X τυχαίο διάνυσμα με $X \sim N_m(\mu, I_m)$ και B είναι ένας $m\times m$ συμμετρικός πίνακας τότε ${}^tXBX \sim \mathcal{X}_k^2(\delta)$ αν και μόνον αν ο B είναι ταυτοδύναμος(δηλ.B^2=B). Σ' αυτή την περίπτωση έχουμε ότι $k = rg(B) = tr(B)$ και $\delta = {}^t\mu B\mu$.

(για την απόδειξη βλέπε [29], Θεώρ.1.4.2, σελ.27)

Λήμμα 5

Έστω $\theta \in \mathbb{R}^p$ και $h : \mathbb{R}^p \to \mathbb{R}$ μετρήσιμη συνάρτηση, $N_p(\theta, I_p)$-ολοκληρώσιμη. Υποθέτουμε ότι η συνάρτηση, από τον $\mathbb{R}^p$ στον $\mathbb{R}^p$, $z \rightsquigarrow h(\|z\|^2)z$ είναι ολοκληρώσιμη (λ^p-ολοκληρώσιμη, όπου λ^p το μέτρο Lebesgue επί του $\mathbb{R}^p$).
Τότε:

$$(i) \int_{\mathbb{R}^p} h(\|z\|^2)(z-\theta)N_p(\theta, I_p; dz) = \frac{\theta}{\|\theta\|^2}\int_{\mathbb{N}_0}\Big(\int_{\mathbb{R}_+} h(u)\mathcal{X}^2_{p+2k}(du)\Big)(2k-\|\theta\|^2)\pi\Big(\frac{\|\theta\|^2}{2}; dk\Big)$$

$$(ii) \int_{\mathbb{R}^p} h(\|z\|^2)zN_p(\theta, I_p; dz) = \frac{\theta}{\|\theta\|^2}\int_{\mathbb{N}_0}\Big(\int_{\mathbb{R}_+} h(u)\mathcal{X}^2_{p+2k}(du)\Big)2k\pi\Big(\frac{\|\theta\|^2}{2}; dk\Big)$$

$$(iii) \int_{\mathbb{R}^p} h(\|z\|^2)zN_p(\theta, I_p; dz) = \theta\int_{\mathbb{N}_0}\Big(\int_{\mathbb{R}_+} h(u)\mathcal{X}^2_{p+2t+2}(du)\Big)\pi\Big(\frac{\|\theta\|^2}{2}; dt\Big)$$

$$(iv) \int_{\mathbb{R}^p} h(\|z\|^2)\,{}^t(z-\theta)zN_p(\theta, I_p; dz) = \int_{\mathbb{N}_0}\Big(\int_{\mathbb{R}_+} h(u)(u-2k)\mathcal{X}^2_{p+2k}(du)\Big)\pi\Big(\frac{\|\theta\|^2}{2}; dk\Big)$$

Απόδειξη

(i) Για κάθε i, με $1 \le i \le p$, έχουμε ότι:

$$\begin{aligned}
\frac{\partial}{\partial\theta_i}\int_{\mathbb{R}^p} h(\|z\|^2)N_p(\theta, I_p; dz) &= \frac{\partial}{\partial\theta_i}\int_{\mathbb{R}^p} h\Big(\sum_{i=1}^p z_i^2\Big)(2\pi)^{-p/2}exp\Big\{-\frac{1}{2}\sum_{i=1}^p(z_i-\theta_i)^2\Big\}\prod_{i=1}^p dz_i \\
&= \int_{\mathbb{R}^p} h\Big(\sum_{i=1}^p z_i^2\Big)(2\pi)^{-p/2}(z_i-\theta_i)exp\Big\{-\frac{1}{2}\sum_{i=1}^p(z_i-\theta_i)^2\Big\}\prod_{i=1}^p dz_i \\
&= \int_{\mathbb{R}^p} h(\|z\|^2)(z_i-\theta_i)N_p(\theta, I_p; dz)
\end{aligned}$$

Όμως, θέτοντας $u{=}\|z\|^2$ και χρησιμοποιώντας τις εκφράσεις (1) και (3) αντίστοιχα του Λήμματος 1.1, έχουμε:

$$\begin{aligned}
\int_{\mathbb{R}^p} h(\|z\|^2)N_p(\theta, I_p; dz) &= \int_{\mathbb{R}_+} h(u)\mathcal{X}^2_p(\|\theta\|^2; du) \\
&= \int_{\mathbb{N}_0} f(k)\pi\Big(\frac{\|\theta\|^2}{2}; dk\Big) \quad , \quad \text{όπου} \quad f(k) = \int_{\mathbb{R}_+} h(u)\mathcal{X}^2_{p+2k}(du) \qquad (3)
\end{aligned}$$

απ᾿ όπου λαμβάνουμε ότι:

$$\begin{aligned}
&\frac{\partial}{\partial\theta_i}\int_{\mathbb{R}^p} h(\|z\|^2)N_p(\theta, I_p; dz) \\
&= \frac{\partial}{\partial\theta_i}\sum_{k=0}^{+\infty} f(k)exp\Big\{-\frac{1}{2}\sum_{i=1}^p\theta_i^2\Big\}\Big(\frac{1}{2}\sum_{i=1}^p\theta_i^2\Big)^k\frac{1}{k!} \\
&= \sum_{k=0}^{+\infty}\frac{f(k)}{k!}\Big[-\theta_i exp\Big\{-\frac{1}{2}\sum_{i=1}^p\theta_i^2\Big\}\Big(\frac{1}{2}\sum_{i=1}^p\theta_i^2\Big)^k + exp\Big\{-\frac{1}{2}\sum_{i=1}^p\theta_i^2\Big\}k\theta_i\Big(\frac{1}{2}\sum_{i=1}^p\theta_i^2\Big)^{k-1}\Big] \\
&= \sum_{k=0}^{+\infty}\frac{f(k)}{k!}\Big[-\theta_i exp\Big\{-\frac{\|\theta\|^2}{2}\Big\}\Big(\frac{\|\theta\|^2}{2}\Big)^k + exp\Big\{-\frac{\|\theta\|^2}{2}\Big\}k\theta_i\Big(\frac{\|\theta\|^2}{2}\Big)^{k-1}\Big]
\end{aligned}$$

$$= \sum_{k=0}^{+\infty} \frac{f(k)}{k!} exp\Big\{-\frac{\|\theta\|^2}{2}\Big\}\Big(\frac{\|\theta\|^2}{2}\Big)^{k-1}\theta_i\Big(k-\frac{\|\theta\|^2}{2}\Big)\Big]$$
$$= \frac{\theta_i}{\|\theta\|^2}\sum_{k=0}^{+\infty} \frac{f(k)}{k!} exp\Big\{-\frac{\|\theta\|^2}{2}\Big\}\Big(\frac{\|\theta\|^2}{2}\Big)^{k}\Big(2k-\|\theta\|^2\Big)\Big]$$
$$= \frac{\theta_i}{\|\theta\|^2}\int_{\mathbb{N}_0} f(k)(2k-\|\theta\|^2)\pi\Big(\frac{\|\theta\|^2}{2};dk\Big)$$

Συνεπώς, προκύπτει ότι για κάθε i με $1{\leq}i{\leq}p$,

$$\int_{\mathbb{R}^p} h(\|z\|^2)(z_i-\theta_i)N_p(\theta,I_p\,;dz) = \frac{\theta_i}{\|\theta\|^2}\int_{\mathbb{N}_0} f(k)(2k-\|\theta\|^2)\pi\Big(\frac{\|\theta\|^2}{2};dk\Big) \tag{4}$$

και επομένως,

$$\int_{\mathbb{R}^p} h(\|z\|^2)(z-\theta)N_p(\theta,I_p\,;dz) = \frac{\theta}{\|\theta\|^2}\int_{\mathbb{N}_0}\Big(\int_{\mathbb{R}_+} h(u)\mathcal{X}^2_{p+2k}(du)\Big)(2k-\|\theta\|^2)\pi\Big(\frac{\|\theta\|^2}{2};dk\Big)$$

(ii) Από τις σχέσεις (3) και (4) προκύπτει ότι, για κάθε i, με $1{\leq}i{\leq}p$,

$$\int_{\mathbb{R}^p} h(\|z\|^2)z_iN_p(\theta,I_p\,;dz) = \frac{\theta_i}{\|\theta\|^2}\int_{\mathbb{N}_0} f(k)2k\pi\Big(\frac{\|\theta\|^2}{2};dk\Big) \tag{5}$$

Συνεπώς,

$$\int_{\mathbb{R}^p} h(\|z\|^2)zN_p(\theta,I_p\,;dz) = \frac{\theta}{\|\theta\|^2}\int_{\mathbb{N}_0}\Big(\int_{\mathbb{R}_+} h(u)\mathcal{X}^2_{p+2k}(du)\Big)2k\pi\Big(\frac{\|\theta\|^2}{2};dk\Big)$$

(iii) Παρατηρούμε ότι:

$$\begin{aligned}
\frac{1}{\|\theta\|^2}\int_{\mathbb{N}_0} f(k)2k\pi\Big(\frac{\|\theta\|^2}{2};dk\Big) &= \frac{2}{\|\theta\|^2}\sum_{k=1}^{+\infty} f(k)k exp\Big\{-\frac{\|\theta\|^2}{2}\Big\}\Big(\frac{\|\theta\|^2}{2}\Big)^{k}\frac{1}{k!}\\
&= \frac{2}{\|\theta\|^2}\sum_{k=1}^{+\infty} f(k) exp\Big\{-\frac{\|\theta\|^2}{2}\Big\}\Big(\frac{\|\theta\|^2}{2}\Big)^{k}\frac{1}{(k-1)!}\\
&= \sum_{k=1}^{+\infty} f(k) exp\Big\{-\frac{\|\theta\|^2}{2}\Big\}\Big(\frac{\|\theta\|^2}{2}\Big)^{k-1}\frac{1}{(k-1)!}\\
(\text{θέτοντας }\ k-1=t) \qquad &= \sum_{t=0}^{+\infty} f(t+1) exp\Big\{-\frac{\|\theta\|^2}{2}\Big\}\Big(\frac{\|\theta\|^2}{2}\Big)^{t}\frac{1}{t!}\\
&= \int_{\mathbb{N}_0} f(t+1)\pi\Big(\frac{\|\theta\|^2}{2};dt\Big)
\end{aligned}$$

Απ' όπου, σύμφωνα με την (ii), συμπεραίνουμε ότι:

$$\int_{\mathbb{R}^p} h(\|z\|^2)z\,N_p(\theta,I_p\,;dz) = \theta\int_{\mathbb{N}_0}\Big(\int_{\mathbb{R}_+} h(u)\mathcal{X}^2_{p+2t+2}(du)\Big)\pi\Big(\frac{\|\theta\|^2}{2};dt\Big)$$

(iv) Τέλος, από την έκφραση (3) του Λήμματος 1 και το (ii) του παρόντος Λήμματος, λαμβάνουμε ότι:

$$\int_{\mathbb{R}^p} h(\|z\|^2)\,{}^t(z-\theta)z\,N_p(\theta, I_p; dz)$$
$$= \int_{\mathbb{R}^p} h(\|z\|^2)\|z\|^2 N_p(\theta, I_p; dz) - {}^t\theta \int_{\mathbb{R}^p} h(\|z\|^2)z\,N_p(\theta, I_p; dz)$$
$$= \int_{\mathbb{R}_+} h(u)u\mathcal{X}_p^2\Big(\|\theta\|^2; du\Big) - {}^t\theta \int_{\mathbb{R}^p} h(\|z\|^2)z\,N_p(\theta, I_p; dz)$$
$$= \int_{\mathbb{N}_0}\Big(\int_{\mathbb{R}_+} h(u)u\mathcal{X}_{p+2k}^2(du)\Big)\pi\Big(\frac{\|\theta\|^2}{2}; dk\Big) - {}^t\theta\frac{\theta}{\|\theta\|^2}\int_{\mathbb{N}_0}\Big(\int_{\mathbb{R}_+} h(u)\mathcal{X}_{p+2k}^2(du)\Big)2k\,\pi\Big(\frac{\|\theta\|^2}{2}; dk\Big)$$
$$= \int_{\mathbb{N}_0}\Big(\int_{\mathbb{R}_+} h(u)(u-2k)\mathcal{X}_{p+2k}^2(du)\Big)\,\pi\Big(\frac{\|\theta\|^2}{2}; dk\Big)$$

$\square$

Λήμμα 6

Έστω $w \sim N(\delta, 1)$ και $\varphi(w^2)$-Borel μετρήσιμη συνάρτηση. Τότε:

$$\mathbb{E}[\varphi(w^2)w^2] = \int_{\mathbb{R}} \varphi(w^2)w^2 N(\delta, 1; dw) = \int_{\mathbb{R}_+} \varphi(u)\mathcal{X}_3^2(\delta^2; du) + \int_{\mathbb{R}_+} \varphi(u)\mathcal{X}_5^2(\delta^2; du)$$

Απόδειξη

Έχουμε:

$$\int_{\mathbb{R}} \varphi(w^2)w^2 N(\delta, 1; dw) = \int_{\mathbb{R}_+} \varphi(u)u\mathcal{X}_1^2(\delta^2; du) = \int_{\mathbb{N}_0}\Big(\int_{\mathbb{R}_+} \varphi(u)u\mathcal{X}_{1+2k}^2(du)\Big)\pi(\delta^2/2; dk)$$

όπου

$$\begin{aligned}
\int_{\mathbb{R}_+} \varphi(u)u\mathcal{X}_{1+2k}^2(du) &= \frac{1}{\Gamma(k+1/2)2^{k+1/2}}\int_0^{+\infty} \varphi(u)uu^{k+\frac{1}{2}-1}e^{-u/2}du \\
&= \frac{1}{\Gamma(k+1/2)2^{k+1/2}}\int_0^{+\infty} \varphi(u)u^{k+\frac{3}{2}-1}e^{-u/2}du \\
&= \frac{\Gamma(k+\frac{3}{2})2^{k+\frac{3}{2}}}{\Gamma(k+1/2)2^{k+1/2}}\int_0^{+\infty} \varphi(u)\mathcal{X}_{2k+3}^2(du) \\
&= \frac{(k+\frac{1}{2})\Gamma(k+\frac{1}{2})2^{k+\frac{3}{2}}}{\Gamma(k+1/2)2^{k+1/2}}\int_0^{+\infty} \varphi(u)\mathcal{X}_{2k+3}^2(du) \\
&= 2(k+1/2)\int_0^{+\infty} \varphi(u)\mathcal{X}_{2k+3}^2(du)
\end{aligned}$$

Συνεπώς,

$$\begin{aligned}
\int_{\mathbb{R}} \varphi(w^2)w^2 N(\delta, 1; dw) &= \int_{\mathbb{N}_0}(2k+1)\Big(\int_{\mathbb{R}_+} \varphi(u)\mathcal{X}_{2k+3}^2(du)\Big)\pi(\delta^2/2; dk) \\
&= 2\int_{\mathbb{N}_0} k\Big(\int_{\mathbb{R}_+} \varphi(u)\mathcal{X}_{2k+3}^2(du)\Big)\pi(\delta^2/2; dk) \\
&\quad + \int_{\mathbb{N}_0}\Big(\int_{\mathbb{R}_+} \varphi(u)\mathcal{X}_{2k+3}^2(du)\Big)\pi(\delta^2/2; dk) \\
&= A + B
\end{aligned}$$

όπου
$$B=\int_{\mathbb{R}_+}\varphi(u)\mathcal{X}_3^2(\delta^2;du)$$
και
$$\begin{aligned}A &= 2\sum_{k=1}^{+\infty}ke^{-\delta^2/2}\frac{(\delta^2/2)^k}{k!}\int_0^{+\infty}\varphi(u)\mathcal{X}_{2k+3}^2(du)\\ &= 2\sum_{k=1}^{+\infty}e^{-\delta^2/2}\frac{(\delta^2/2)^k}{(k-1)!}\int_0^{+\infty}\varphi(u)\mathcal{X}_{2k+3}^2(du)\\ &= 2\frac{\delta^2}{2}\sum_{m=0}^{+\infty}e^{-\delta^2/2}\frac{(\delta^2/2)^m}{m!}\int_0^{+\infty}\varphi(u)\mathcal{X}_{2m+5}^2(du)\\ &= \delta^2\int_{\mathbb{R}_+}\varphi(u)\mathcal{X}_5^2(\delta^2;du)\end{aligned}$$
απ' όπου προκύπτει το ζητούμενο.

□

Λήμμα 7

Υποθέτουμε ότι $w\sim N(\delta,1)$ και $\varphi(w^2)$-Borel μετρήσιμη συνάρτηση. Τότε:
$$\mathbb{E}[\varphi(w^2)w]=\int_{\mathbb{R}}\varphi(w^2)wN(\delta,1;dw)=\delta\int_{\mathbb{R}_+}\varphi(u)\mathcal{X}_3^2(\delta^2;du)$$

Απόδειξη

Υπό τις υποθέσεις εκείνες οι οποίες επιτρέπουν την εναλλαγή παραγώγου και ολοκληρώματος έχουμε:
$$\begin{aligned}&\int_{\mathbb{R}}\varphi(w^2)wN(\delta,1;dw)\\ &=\frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}}\varphi(w^2)wexp\Big\{-\frac{1}{2}(w-\delta)^2\Big\}dw\\ &=\frac{1}{\sqrt{2\pi}}exp\Big\{-\frac{\delta^2}{2}\Big\}\int_{\mathbb{R}}\varphi(w^2)wexp\Big\{-\frac{1}{2}w^2+\delta w\Big\}dw\\ &=\frac{1}{\sqrt{2\pi}}exp\Big\{-\frac{\delta^2}{2}\Big\}\int_{\mathbb{R}}\varphi(w^2)\frac{\partial}{\partial\delta}exp\Big\{-\frac{1}{2}w^2+\delta w\Big\}dw\\ &=\frac{1}{\sqrt{2\pi}}exp\Big\{-\frac{\delta^2}{2}\Big\}\frac{\partial}{\partial\delta}\int_{\mathbb{R}}\varphi(w^2)exp\Big\{-\frac{1}{2}w^2+\delta w\Big\}dw\\ &=exp\Big\{-\frac{\delta^2}{2}\Big\}\frac{\partial}{\partial\delta}\Bigg(exp\Big\{\frac{\delta^2}{2}\Big\}\int_{\mathbb{R}}\varphi(w^2)N(\delta,1;dw)\Bigg)\\ &=exp\Big\{-\frac{\delta^2}{2}\Big\}\frac{\partial}{\partial\delta}\Bigg(exp\Big\{\frac{\delta^2}{2}\Big\}\int_{\mathbb{R}_+}\varphi(w)\mathcal{X}_1^2(\delta^2;du)\Bigg)\\ &=exp\Big\{-\frac{\delta^2}{2}\Big\}\frac{\partial}{\partial\delta}\Bigg(exp\Big\{\frac{\delta^2}{2}\Big\}\sum_{k=0}^{+\infty}e^{-\delta^2/2}\frac{(\delta^2/2)^k}{k!}\Big(\int_{\mathbb{R}_+}\varphi(u)\mathcal{X}_{1+2k}^2(du)\Big)\Bigg)\end{aligned}$$

$$= exp\Big\{-\frac{\delta^2}{2}\Big\}\frac{\partial}{\partial\delta}\sum_{k=0}^{+\infty}\frac{(\delta^2/2)^k}{k!}\int_{\mathbb{R}_+}\varphi(u)\mathcal{X}^2_{1+2k}(du)$$

$$= exp\Big\{-\frac{\delta^2}{2}\Big\}\sum_{k=0}^{+\infty}\frac{k(\delta^2/2)^{k-1}}{k!}\delta\int_{\mathbb{R}_+}\varphi(u)\mathcal{X}^2_{1+2k}(du)$$

$$= exp\Big\{-\frac{\delta^2}{2}\Big\}\sum_{k=1}^{+\infty}\frac{1}{(k-1)!}(\delta^2/2)^{k-1}\delta\int_{\mathbb{R}_+}\varphi(u)\mathcal{X}^2_{1+2k}(du)$$

$$= \delta exp\Big\{-\frac{\delta^2}{2}\Big\}\sum_{m=0}^{+\infty}\frac{1}{m!}(\delta^2/2)^m\int_{\mathbb{R}_+}\varphi(u)\mathcal{X}^2_{2m+3}(du)$$

$$= \delta\int_{\mathbb{R}_+}\varphi(u)\mathcal{X}^2_3(\delta^2;du)$$

□

Θεώρημα 1

Αν $w \sim N_k(\theta, I_k)$ *και* $\varphi(w^2)$*-Borel μετρήσιμη συνάρτηση, τότε:*

$$\begin{aligned}\mathbb{E}[\varphi({}^tww)w\,{}^tw] &= \int_{\mathbb{R}^k}\varphi({}^tww)w\,{}^twN_k(\theta, I_k; dw)\\ &= \Big(\int_{\mathbb{R}_+}\varphi(u)\mathcal{X}^2_{k+2}(2\lambda;du)\Big)I_k + \Big(\int_{\mathbb{R}_+}\varphi(u)\mathcal{X}^2_{k+4}(2\lambda;du)\Big)\theta\,{}^t\theta\end{aligned}$$

όπου $\lambda = {}^t\theta\theta/2$.

Απόδειξη

Έστω $w = {}^t(w_1, ..., w_k)$ και καθορίζουμε τα διαγώνια και μη-διαγώνια στοιχεία του πίνακα $\mathbb{E}[\varphi({}^tww)w\,{}^tw]$.

Τα διαγώνια στοιχεία του είναι της μορφής :

$$\mathbb{E}\Big[\varphi\Big(\sum_{i=1}^k w_i^2\Big)w_i^2\Big]$$

τα οποία γράφονται:

$$\begin{aligned}\mathbb{E}\Big[\varphi\Big(\sum_{i=1}^k w_i^2\Big)w_i^2\Big] &= \int_{\mathbb{R}^k}\varphi\Big(w_i^2+\sum_{j\neq i}w_j^2\Big)w_i^2N_k(\theta, I_k; dw)\\ &= \int_{\mathbb{R}^{k-1}}\Big(\int_{\mathbb{R}}\varphi\Big(w_i^2+\sum_{j\neq i}w_j^2\Big)w_i^2N(\theta_i, 1; dw_i)\Big)N_{k-1}(\theta^i, I_{k-1}; dw^i)\end{aligned}$$

όπου $\theta^i = (\theta_1, ..., \theta_{i-1}, \theta_{i+1}, ..., \theta_k)$ και $w^i = (w_1, ..., w_{i-1}, w_{i+1}, ..., w_k)$.

Επομένως, λόγω του Λήμματος 6 (ανωτέρω), έχουμε:

$$
\begin{aligned}
&\mathbb{E}\Big[\varphi\Big(\sum_{i=1}^{k} w_i^2\Big) w_i^2\Big] \\
&= \int_{\mathbb{R}^{k-1}} \bigg(\int_{\mathbb{R}_+} \varphi\Big(u + \sum_{j\neq i} w_j^2\Big)\mathcal{X}_3^2(\theta_i^2; du) + \theta_i^2 \int_{\mathbb{R}_+} \varphi\Big(u + \sum_{j\neq i} w_j^2\Big)\mathcal{X}_5^2(\theta_i^2; du)\bigg) N_{k-1}(\theta^i, I_{k-1}; dw^i) \\
&= \int_{\mathbb{R}_+} \bigg(\int_{\mathbb{R}^{k-1}} \varphi\Big(u + \sum_{j\neq i} w_j^2\Big) N_{k-1}(\theta^i, I_{k-1}; dw^i)\bigg)\mathcal{X}_3^2(\theta_i^2; du) \\
&\quad + \theta_i^2 \int_{\mathbb{R}_+} \bigg(\int_{\mathbb{R}^{k-1}} \varphi\Big(u + \sum_{j\neq i} w_j^2\Big) N_{k-1}(\theta^i, I_{k-1}; dw^i)\bigg)\mathcal{X}_5^2(\theta_i^2; du) \\
&= \int_{\mathbb{R}_+} \bigg(\int_{\mathbb{R}_+} \varphi(u+v)\mathcal{X}_{k-1}^2({}^t\theta^i\theta^i; dv)\bigg)\mathcal{X}_3^2(\theta_i^2; du) \\
&\quad + \theta_i^2 \int_{\mathbb{R}_+} \bigg(\int_{\mathbb{R}_+} \varphi(u+v)\mathcal{X}_{k-1}^2({}^t\theta^i\theta^i; dv)\bigg)\mathcal{X}_5^2(\theta_i^2; du) \\
&= \int_{\mathbb{R}_+\times\mathbb{R}_+} \varphi(u+v)\mathcal{X}_{k-1}^2({}^t\theta^i\theta^i; dv)\mathcal{X}_3^2(\theta_i^2; du) + \theta_i^2 \int_{\mathbb{R}_+\times\mathbb{R}_+} \varphi(u+v)\mathcal{X}_{k-1}^2({}^t\theta^i\theta^i; dv)\mathcal{X}_5^2(\theta_i^2; du) \\
&= \int_{\mathbb{R}_+} \varphi(t)\mathcal{X}_{k+2}^2({}^t\theta\theta; dt) + \theta_i^2 \int_{\mathbb{R}_+} \varphi(t)\mathcal{X}_{k+4}^2({}^t\theta\theta; dt)
\end{aligned}
$$

η τελευταία αυτή ισότητα προκύπτει από την ανεξαρτησία των u και v και το Λήμμα 3.

Τα μη-διαγώνια στοιχεία του, για $i \neq j$, έχουν τη μορφή:

$$
\mathbb{E}\Big[\varphi\Big(\sum_{i=1}^{k} w_i^2\Big) w_i w_j\Big] = \int_{\mathbb{R}^{k-1}} w_j \bigg(\int_{\mathbb{R}} \varphi\Big(w_i^2 + \sum_{l\neq i} w_l^2\Big) w_i N(\theta_i, 1; dw_i)\bigg) N_{k-1}(\theta^i, I_{k-1}; dw^i)
$$

τα οποία λόγω του Λήμματος 7 γράφονται:

$$
\begin{aligned}
\mathbb{E}\Big[\varphi\Big(\sum_{i=1}^{k} w_i^2\Big) w_i w_j\Big] &= \int_{\mathbb{R}^{k-1}} w_j \bigg(\theta_i \int_{\mathbb{R}_+} \varphi\Big(u + \sum_{l\neq i} w_l^2\Big)\mathcal{X}_3^2(\theta_i^2; du)\bigg) N_{k-1}(\theta^i, I_{k-1}; dw^i) \\
&= \theta_i \int_{\mathbb{R}_+} \bigg(\int_{\mathbb{R}^{k-1}} \varphi\Big(u + \sum_{l\neq i} w_l^2\Big) w_j N_{k-1}(\theta^i, I_{k-1}; dw^i)\bigg)\mathcal{X}_3^2(\theta_i^2; du)
\end{aligned}
$$

Εφαρμόζοντας ακόμη μία φορά το Λήμμα 7 στο παραπάνω εσωτερικό ολοκλήρωμα και θέτοντας $\theta^{i,j} = (\theta_1, ..., \theta_{i-1}, \theta_{i+1}, ..., \theta_{j-1}, \theta_{j+1}, ..., \theta_k)$, λαμβάνουμε:

$$
\begin{aligned}
&\int_{\mathbb{R}^{k-1}} \varphi\Big(u + \sum_{l\neq i} w_l^2\Big) w_j N_{k-1}(\theta^i, I_{k-1}; dw^i) \\
&= \int_{\mathbb{R}^{k-1}} \varphi\Big(u + w_j^2 + \sum_{l\neq i,j} w_l^2\Big) w_j N_{k-1}(\theta^i, I_{k-1}; dw^i) \\
&= \int_{\mathbb{R}^{k-2}} \bigg(\int_{\mathbb{R}} \varphi\Big(w_j^2 + u + \sum_{l\neq i,j} w_l^2\Big) w_j N(\theta_j, 1; dw_j)\bigg) N_{k-2}(\theta^{i,j}, I_{k-2}; dw^{i,j}) \\
&= \int_{\mathbb{R}^{k-2}} \bigg(\theta_j \int_{\mathbb{R}_+} \varphi\Big(v + u + \sum_{l\neq i,j} w_l^2\Big)\mathcal{X}_3^2(\theta_j^2; dv)\bigg) N_{k-2}(\theta^{i,j}, I_{k-2}; dw^{i,j})
\end{aligned}
$$

$$= \theta_j \int_{\mathbb{R}_+} \left(\int_{\mathbb{R}^{k-2}} \varphi\Big(v + u + \sum_{l \neq i,j} w_l^2\Big) N_{k-2}(\theta^{i,j}, I_{k-2}; dw^{i,j}) \right) \mathcal{X}_3^2(\theta_j^2; dv)$$
$$= \theta_j \int_{\mathbb{R}_+} \left(\int_{\mathbb{R}_+} \varphi(v + u + z) \mathcal{X}_{k-2}^2\Big(\sum_{l \neq i,j} \theta_l^2; dz\Big) \right) \mathcal{X}_3^2(\theta_j^2; dv)$$

Συνεπώς,

$$\begin{aligned}
\mathbb{E}\Big[\varphi\Big(\sum_{i=1}^k w_i^2\Big) w_i w_j\Big] &= \theta_i \theta_j \int_{\mathbb{R}_+} \left(\int_{\mathbb{R}_+} \left(\int_{\mathbb{R}_+} \varphi(v + u + z) \mathcal{X}_{k-2}^2\Big(\sum_{l \neq i,j} \theta_l^2; dz\Big) \right) \mathcal{X}_3^2(\theta_j^2; dv) \right) \mathcal{X}_3^2(\theta_i^2; du) \\
&= \theta_i \theta_j \int_{\mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R}_+} \varphi(u + v + z) \mathcal{X}_3^2(\theta_i^2; du) \mathcal{X}_3^2(\theta_j^2; dv) \mathcal{X}_{k-2}^2\Big(\sum_{l \neq i,j} \theta_l^2; dz\Big) \\
&= \theta_i \theta_j \int_{\mathbb{R}_+} \varphi(t) \mathcal{X}_{k+4}^2({}^t\theta\theta; dt)
\end{aligned}$$

λόγω ανεξαρτησίας των u, v και z και του Λήμματος 3.

Συνδυάζοντας τα διαγώνια και μη διαγώνια στοιχεία, καταλήγουμε στη ζητούμενη ισότητα πινάκων.

□

Βιβλιογραφία

[1] **Abramowitz M.**, **Stegun I.**(**1964**) Handbook of Mathematical Functions and Formulas, Graphs and Mathematical Tables $-Dover$

[2] **Alam K.**(**1973**) A Family of Admissible Minimax Estimators of the Mean of a Multivariate Normal Distribution $-Annals\ of\ Statistics, 1, pp.517-525$

[3] **Baranchik A.**(**1964**) Multiple Regression and Estimation of the Mean Vector of a Multivariate Normal Distribution $-Stanford\ Univ.$, $Technical\ Report\ N^o$ 51

[4] **Berger J.**(**1976**) Minimax Estimation of a Multivariate Normal Mean for Arbitrary Quadratic Loss $-J.Mult.Anal.$, 6, $pp.256-264$

[5] **Berger J.**(**1978**) Minimax Estimation of a Multivariate Normal Mean under Polynomial Quadratic Loss $-J.Mult.Anal.$, 8, $pp.173-180$

[6] **Berger J.**(**1980**) A Robust Generalised Bayes Estimator and Confidencial Regions for a Multivariate Normal Mean $-Ann.Stat.$, 8(4), $pp.716-761$

[7] **Berger J.**(**1982**) Selecting a Minimax Estimator of a Multivariate Normal Mean $-Ann.$ $Stat.$, 10(1), $pp.81-92$

[8] **Berger J.**, **Berliner L.**(**1984**) Bayesian imput in Stein Estimation and a new Minimax Empirical Estimator $-J.Econometrics$, 25, $pp.87-108$

[9] **Berger J.**, **Bock M.E**(**1976**) Eliminating Singularities of Stein-Type Estimators of Location Vectors $-JRSS\ B$ 37(2), $pp.166-170$

[10] **Berger J.**, **Srinivasan C.**(**1978**) Generalised Bayes Estimators in Multivariate Problems $-Ann.Statist.$, 6(4), $pp.783-801$

[11] **Bhattacharya P.**(**1966**) Estimating the Mean of a Multivariate Normal Population with General Quadratic Loss Function $-Ann.Math.Stat.$, 37, $pp.1819-1825$

[12] **Bock M.E.**(**1975**) Minimax Estimators of the Mean of a Multivariate Normal Distribution $-Ann.Stat.$, 3(1), $pp.209-218$

[13] **Bock M.E.**(**1987**) Shrinkage Estimators : prendo Bayes rules for normal mean vectors $-Techn.Rep.$, $87-18$, $Purdue\ University$

[14] **Brandwein A.C.**, **Strawderman W.**(**1980**) Minimax Estimation of Location Parameters for Spherically Symmetric Distributions with Concave Loss $-Ann.Stat.$, 8(2), $pp.279-284$

[15] **Brandwein A.C.**, **Strawderman W.**(**1991**) Improved Estimates of Location in the Presence of an Unknown Scale $-Journal\ of\ Multivariate\ Analysis$, 39, $pp.305-314$

[16] **Brown L.(1966)** On the Admissibility of Invariant Estimators of one or more Location Parameters $-Ann.Math.Stat.$ 37(5), $pp.1087-1136$

[17] **Brown L.(1971)** Admissible Estimators, Reccurent Diffusions and Insoluble Boundary Problems $-Ann.Math.Statist.$, 42(3), $pp.855-903$

[18] **Brown L.(1975)** Estimation with Incompletely Specified Loss (the case of several location parametres) $-Journal\ of\ the\ American\ Statistical\ Assosiation$, $Vol.70$, N^o 350, $pp.417-427$

[19] **Brown L.(1986)** Fundamentals of Statistical Exponential Families $-Monograph\ Ser.$, $Volume$ 9

[20] **Brown L.(1987)** The Differential Inequality of a Statistical Estimation Problem $-Proc.\ Fourth\ Purdue\ Symp.\ Stat.\ Dec.\ Theo.\ Rel.\ Topics$ ($S.S.Gupta$, $and\ J.Berger$, eds), $Springer\ Verlang$, $New\ York$

[21] **Cellier D. et Fourdrinier D.(1985)** Estimateurs à rétrécisseurs de la moyene d′une loi normale multidimensionelle, pour un coût quadratique général $-Statistique\ et\ Analyse\ des\ Donées$, vol 10(3), $pp.26-41$

[22] **Cellier D. et Fourdrinier D.(1988)** Sur les lois à symmetrie elliptique $-Lecture\ Notes\ in\ Mathematics$

[23] **Cellier D., Fourdrinier D. et Robert C.(1989)** Controlled Shrinkage Estimators(A class of estimators better than the least squares estimator, with respect to a general quadratic loss, for normal observations) $-Statistics$ 20, 1, $pp.13-22$

[24] **Cellier D. et Fourdrinier D.(1994)** Shrinkage Estimators under Spherical Symmetry for the General Linear Model $-$ *Document* $1994-09$, $Université\ de\ Rouen$

[25] **Criticou D., Terzakis D. (1992)** Une Classe d'Estimateurs à Rétrécisseur Bayesiens pour la Moyenne d'une Vecteur Normal $-Statistique\ et\ Analyse\ des\ Donées$, $Vol.16$, N^o 3, $pp.1-23$

[26] **Dey D., Berger J.(1983)** On Truncation of Shrinkage Estimators in Simultaneous Estimation of Normal Means $-JASA$, 78, $pp.865-869$

[27] **Efron B. and Morris C.(1975)** Data Analysis using Stein′s Estimator and its Generalizations $-Journal\ of\ the\ American\ Statistical\ Association$, vol 70, 350, $pp.311-319$

[28] **Efron B. and Morris C.(1976)** Families of Minimax Estimators of the Mean of a Multivariate Normal Distribution $-Ann.Stat.$, 4(1), $pp.11-21$

[29] **Egerton M.F. and Laycock P.J..(1982)** An Explicit Formula for the Risk of James-Stein Estimators $-The\ Canadian\ Journal\ of\ Statistics$, $Vol.10$, N^o 3, $pp.199-205$

[30] **Farrell R.(1968)** On a Necessary and Sufficient Condition for Admissibility of Estimators when Convex Loss is used $-$*Ann.Math.Statist.*, 49 (1), *pp.*$23-28$

[31] **Fourgeaud C. and Fuchs A.(1972)** Statistique $-$*Dunod*

[32] **Fraisse A.M., Robert C. et Roy M.(1987)** Estimateurs à Rétrécisseur Matriciel Différentiable, pour un Coût Quadratique Général $-$*Annales d'Économie et de Statistique*, N^o 8, *pp.*$161-175$

[33] **Fraisse A.M., Raoult J.P., Robert C. et Roy M.(1990)** Une Condition Nécessaire d' Admissibilité et ses Conséquences sur les Estimateurs à Rétrécisseur de la Moyenne d' un Vecteur Normal $-$*The Canadian Journal of Statistics*, *Vol.*18, N^o 3, *pp.*$213-220$

[34] **James W., Stein C.(1961)** Estimation with Quadratic Loss $-$*Proc.4th Berkeley Symp. Math.Stat.Prob.* 1, *pp.*$136-179$

[35] **Judge G. and Bock E.M(1978)** The Statistical Implications of Pre-test and Stein-rule estimators in Econometrics $-$*North-Holland*

[36] **Lehman E.L.** Testing Statistical-Hypotheses

[37] **Manjogue S. and Rao F.(1985)** Simultaneous Estimation of Parametrs upon Compound Weighted Loss $-$*SankhyaB*, 47(3), *pp.*$338-345$

[38] **Moore T. and Brook R.(1978)** Risk Estimate optimality of James-Stein Estimators $-$*Ann.Stat.*, 6(4), *pp.*$917-919$

[39] **Muirhead R.J.(1982)** Aspects of Multivariate Statistical Theory $-$*Wiley Series in Probability and Mathematical Statistics*

[40] **Robert C.(1987)** An Explicit Formula for the Risk of the Positive-Part James-Stein Estimator $-$*Purdue University*, *Technical Report*

[41] **Robert C.(1987)** Resultats Nouveaux sur les Estimateurs à Rétrécisseurs Scalaires et Matriciels $-$*These*, *Université de Rouen*

[42] **Stein C.(1955)** A Necessary and Sufficient Condition for Admissibility $-$*Ann.Math.Stat.*, 26, *pp.*$518-522$

[43] **Stein C.(1956)** Inadmissibility of the Usual Estimator of the Mean of a Mutlivariate Normal Distribution $-$*Third Berkeley Symp.Math.Statist.Probab.*, 1, *pp.*$197-206$.*Univ. of California Press*, *Berkerley*

[44] **Stein C.(1973)** Estimation of the Mean of a Multivariate Normal Distribution $-$*Proc. Prague Symp.Asymp.Stat.*, 1, *pp.*$345-381$

[45] **Stein C.(1981)** Estimation of the Mean of a Multivariate Normal Distribution $-$*The*

Annals of Statistics, vol 9, N^o 6, pp.1135 − 1151

[46] **Strawderman W.**(**1973**) Proper Bayes Minimax Estimators of the Multivariate Normal Mean Vector for the Case of Common Unknown Variances −*Annals of Statistics*, 1(6), pp.1189 − 1194

[47] **Tse Fen Li**, **Wen Hou Kuo**(**1982**) Generalized James-Stein Estimators −*Comm.Stat.*, 11(20), pp.2249 − 2257