«ΣΠΟΥΔΑΙ», Τόμος 41, Τεύχος 2ο, Πανεπιστήμιο Πειραιώς
«SPOUDAI», Vol. 41, No 2, University of Piraeus

ΜΙΑ MONTE CARLO ΜΕΛΕΤΗ ΤΩΝ ΕΚΤΙΜΗΤΩΝ RIDGE ΚΑΙ ΕΛΑΧΙΣΤΩΝ ΤΕΤΡΑΓΩΝΩΝ

Του
Πάνου Αναστ. Πανόπουλου
Οικονομικό Πανεπιστήμιο Αθηνών

Abstract

Multicollinearity is a very severe problem in many statistical and econometrics application.

The most promising technique for reducing the harmful effect of multicillinearity is called Ridge Regression (RR).

At this study ("A MONTE CARLO STUDY OF RIDGE AND OLS ESTIMATORS") a Monte-Carlo experiment was used for the OLS estimator and Ridge estimators which were determined by five different methods of estimations of "k". These estimators were compered using the mean squared error (MSE) and the effectiveness index (EFF) criteria. The Monte-Carlo experiment was based on two models from the Greek economy.

1. Εισαγωγή

Έστω το γραμμικό υπόδειγμα

$$y = X\beta + u, \tag{1}$$

όπου y διάνυσμα τάξεως (nx1) των τιμών της εξαρτημένης μεταβλητής, X μήτρα τάξεως (nxp) των τιμών των ερμηνευτικών (ανεξαρτήτων) μεταβλητών, β διάνυσμα τάξεως (px1) των προς εκτίμηση συντελεστών και u διάνυσμα τάξεως (nx1) των τυχαίων σφαλμάτων.

Η εκτιμήτρια ελαχίστων τετραγώνων $\hat{\beta}$ του διανύσματος β δίνεται από τη λύση του συστήματος των κανονικών εξισώσεων

$$X'X\hat{\beta} = X'y, \tag{2}$$

όπου X′ η ανάστροφη μήτρα της X.

Για να υπάρχει λύση στο σύστημα (2) πρέπει να είναι δυνατή η αντιστροφή της μήτρας $X'X$, δηλαδή πρέπει η ορίζουσα $|X'X|$ να είναι διάφορος του μηδενός. Όταν όμως συμβαίνει να είναι $|X'X|=0$ τότε δεν είναι δυνατή η εκτίμηση του υποδείγματος με τη μέθοδο των ελαχίστων τετραγώνων (OLS). Στην περίπτωση αυτή λέμε ότι έχουμε πλήρη πολυσυγγραμικότητα, ενώ όταν συμβαίνει η τιμή της $|X'X|$ να είναι κοντά στο μηδέν λέμε ότι έχουμε μη πλήρη πολυσυγγραμικότητα. Στη δεύτερη περίπτωση η εκτίμηση του υποδείγματος είναι δυνατή, είναι όμως αδύνατο να προσδιορισθούν οι ακριβείς τιμές των επί μέρους παραμέτρων.

Μία μέθοδος θεραπείας του προβλήματος της πολυσυγγραμικότητας είναι η Ridge Regression, όπου στη μέθοδο εκτιμήσεως (εκτιμήτρια Ridge) υπεισέρχεται μία σταθερά k (Hoerl και Kennartd 1970 a,b). Μία από τις βασικές δυσκολίες στη μέθοδο αυτή είναι ο προσδιορισμός της σταθεράς k. Στη βιβλιογραφία έχει προταθεί μεγάλος αριθμός μεθόδων προσδιορισμού του k. Στην εργασία αυτή με την τεχνική Monte – Carlo γίνεται σύγκριση διαφόρων μεθόδων προσδιορισμού των τιμών του k καθώς και των αντίστοιχων εκτιμητριών Ridge με βάση αυτές τις τιμές του k. Για τη σύγκριση αυτή χρησιμοποιήθηκαν υποδείγματα με δύο και τρεις ερμηνευτικές μεταβλητές και δεδομένα από την Ελληνική Οικονομία.

Η διάρθρωση της εργασίας αυτής έχει ως εξής: Στην παράγραφο 2 γίνεται περιγραφή μεθόδων Ridge Regression καθώς και τεσσάρων μεθόδων προσδιορισμού της σταθεράς k. Στην παράγραφο 3 αναπτύσσεται η μέθοδος προσδιορισμού του k η οποία προτάθηκε από τον Lee (1979). Εμπειρική αξιολόγηση όλων των μεθόδων προσδιορισμού του k γίνεται στην παράγραφο 4. Επίσης στην παράγραφο 5 με την τεχνική Monte – Carlo γίνεται σύγκριση τόσο των μεθόδων προσδιορισμού του k όσο και των μέσων τετραγωνικών σφαλμάτων (MSE) των εκτιμητριών που επιτυγχάνονται μ' αυτές τις τιμές του k. Τέλος στη παράγραφο 6 παρουσιάζονται τα γενικά συμπεράσματα.

2. Ridge regression

Αν στο υπόδειγμα (1) η μήτρα X και το διάνυσμα y τυποποιηθούν τότε η μήτρα $X'X$ είναι η μήτρα συσχετίσεως και το $X'y$ είναι το διάνυσμα των συντελεστών απλής συσχετίσεως. Στην περίπτωση αυτή η εκτιμήτρια ελαχίστων τετραγώνων $\hat{\beta}$ του διανύσματος β είναι

$$\hat{\beta} = (X'X)^{-1}X'y. \tag{3}$$

Επίσης αν $\lambda_{min} = \lambda_1 \leqslant \lambda_2 \leqslant .... \leqslant \lambda_p = \lambda_{max}$
είναι οι p χαρακτηριστικές ρίζες της X΄X τότε το μέσο τετραγωνικό σφάλμα (MSE) της $\hat{\beta}$ γράφεται

$$MSE\ (\hat{\beta}) = \sigma^2 tr(X'X)^{-1} = \sigma^2 \sum_{i=1}^{p} \lambda i^{-1}. \tag{4}$$

Ακόμη αν η μήτρα X΄X είναι η μοναδιαία μήτρα, δηλαδή όταν οι ερμηνευτικές μεταβλητές είναι ορθογώνιες («ιδανική περίπτωση»), τότε $\lambda_1 = \lambda_2 = ... = \lambda_p = 1$ και η (4) γράφεται

$$MSE\ (\hat{\beta}) = p\sigma^2 \tag{5}$$

Η διαπίστωση και η μέτρηση της πολυσυγγραμικότητας είναι ένα από τα σοβαρότερα προβλήματα που αντιμετωπίζονται στη Στατιστική και την Οικονομετρία. Στη βιβλιογραφία έχει προταθεί μεγάλος αριθμός δεικτών και κριτηρίων μετρήσεως της πολυσυγγραμικότητας, οι χρησιμοποιούμενοι στην παρούσα εργασία δίνονται στον πίνακα 1.

Πίνακας 1

Δείκτες και κριτήρια μετρήσεως της πολυσυγγραμμικότητας

	Τύπος	Φράγματα	Ιδανική τιμή	Παρατηρήσεις
1	$M = \sum_{i=1}^{p} \lambda i^{-1}$	$M \geqslant p$	$M = p$	Mahayan et. al. (1977), Lawless (1978)
2	maxVIF	maxVIF$\geqslant$1	maxVIF = 1	Marquardt and Snee (1975)
3	$maxR^2$	$0 \leqslant maxR^2 \leqslant 1$	$maxR^2 = 0$	Kmenta (1971)

Στην προσπάθεια να ελαττωθεί το πρόβλημα της πολυσυγγραμικότητας οι Hoerl και Kennartd (1970 a,b) ανέπτυξαν μια πολλά υποσχόμενη μέθοδο την οποία ονόμασαν Rigde Regression (RR). Η μέθοδος αυτή στηρίζεται στο κριτήριο ελαχιστοποιήσεως του μέσου τετραγωνικού σφάλματος (MSE). Η RR είναι η μέθοδος όπου σε όλα τα στοιχεία της κυρίας διαγωνίου της μήτρας συσχετίσεως X΄X προστίθεται η αυτή θετική σταθερά k. Η εκτιμήτρια Ridge $\hat{\beta}^*$ του διανύσματος β δίνεται από τη σχέση

$$\hat{\beta}^* = (X'X + Ik)^{-1}X'y, \qquad k > 0 \tag{6}$$

και είναι μεροληπτική.

Το μέσο τετραγωνικό σφάλμα της $\hat{\beta}^*$ δίνεται από τον τύπο

$$\text{MSE}\ (\hat{\beta}^*) = \sigma^2 \sum_{i=1}^{p} \frac{\lambda_i}{(\lambda i + k)^2} + k^2 \sum_{i=1}^{p} \frac{\alpha i^2}{(\lambda i + k)^2}, \tag{7}$$

όπου $\alpha = V'\beta$ και V ορθογώνια μήτρα με στήλες τα ιδιοδιανύσματα της μήτρας συσχετίσεως $X'X$ για τις χαρακτηριστικές ρίζες της $\lambda_1, \lambda_2, \ldots, \lambda_p$.

Αποδεικνύεται ότι οι εκτιμήτριες $\hat{\alpha}$ και $\hat{\beta}$ συνδέονται μεταξύ τους με τη σχέση

$$\hat{\alpha} = V'\hat{\beta}. \tag{8}$$

Από την (7) προκύπτει ότι το MSE $(\hat{\beta}^*)$ δεν μπορεί να υπολογισθεί αφού είναι συνάρτηση των άγνωστων παραμέτρων σ^2 και α. Αποδεικνύεται επίσης ότι υπάρχει πάντοτε $k > 0$. (Hoerl and Kennard (1970)) ώστε να ισχύει

$$\text{MSE}(\hat{\beta}^*) < \text{MSE}\ (\hat{\beta}). \tag{9}$$

Ένα από τα μειονεκτήματα της μεθόδου RR είναι όπως αναφέρθηκε ότι το MSE $(\hat{\beta}^*)$ είναι συνάρτηση των άγνωστων παραμέτρων β και σ^2, οπότε δεν μπορεί να προσδιορισθεί απευθείας η τιμή του k. Στη βιβλιογραφία έχει προταθεί μεγάλος αριθμός μεθόδων για τον προσδιορισμό του k, αυτές που θα χρησιμοποιηθούν στην παρούσα εργασία είναι οι εξής:

(α) Μέθοδος HKB (Hoerl, Kennard and Baldwin (1975)), όπου η τιμή του k υπολογίζεται από τη σχέση

$$k = \frac{p\hat{\sigma}^2}{\hat{\beta}'\hat{\beta}}. \tag{10}$$

(β) Μέθοδος HSL (Hocking, Speed and Lynn (1976)), όπου η τιμή του k προσδιορίζεται από τη σχέση

$$k = \frac{\hat{\sigma}^2 \sum_{i=1}^{p} \lambda_i^2 \hat{\alpha}_i^2}{\sum_{i=1}^{p} \lambda_i^2 \hat{\alpha}_i^4}. \tag{11}$$

(γ) Μέθοδος DSW (Dempster, Schatzoft and Wermuth (1977)), όπου η τιμή του k προσδιορίζεται από τη σχέση

$$p = \sum_{i=1}^{p} \frac{\hat{\alpha}_i^2}{\hat{\sigma}^2 (1/k + 1/\lambda_i)} . \tag{12}$$

(δ) Μέθοδος LW (Lawless and Wang (1976)), όπου η τιμή του k προσδιορίζεται από τη σχέση

$$k = \frac{p\hat{\sigma}^2}{\sum_{i=1}^{p} \lambda_i \hat{\alpha}_i^2} . \tag{13}$$

Πρέπει να σημειωθεί ότι όλες οι προαναφερθείσες μέθοδοι υπολογισμού του k εκφράζονται ως συνάρτηση των $\hat{\sigma}^2$ και $\hat{\alpha}_i$ (ή$\hat{\beta}_i$) που εκτιμήθηκαν από το αρχικό υπόδειγμα στο οποίο παρουσιάζεται το πρόβλημα της πολυσυγγραμμικότητας.

3. Προσδιορισμός του k με τη μέθοδο VNC [Lee (1979)]

Στην παράγραφο αυτή θα αναπτυχθεί η μέθοδος προσδιορισμού του k η οποία προτάθηκε από τον Lee (1979) και ονομάστηκε VNC (Variance Normalization Criterion). Η προσπάθεια της μεθόδου αυτής είναι η αποφυγή των μειονεκτημάτων των μεθόδων προσδιορισμού της παραμέτρου k.

Ο Lee προτείνει τον απευθείας προσδιορισμό της σταθεράς k από τη σχέση

$$\sum_{i=1}^{p} \frac{\lambda_i}{(\lambda_i + k)^2} = p. \tag{14}$$

Η φιλοσοφία αυτής της μεθόδου βρίσκεται στην παρατήρηση ότι για την «ιδανική περίπτωση» στην μέθοδο OLS είναι σύμφωνα με τις (4) και (5)

$$\sigma^2 \sum_{i=1}^{p} \lambda i^{-1} = p\sigma^2, \tag{15}$$

τότε, για την «ιδανική περίπτωση» στη μέθοδο RR θα είναι σύμφωνα με την (7)

$$\sigma^2 \sum_{i=1}^{p} \frac{\lambda_i}{(\lambda_i + k)^2} = p\sigma^2, \tag{16}$$

οπότε προκύπτει ότι

$$\sum_{i=1}^{p} \frac{\lambda_i}{(\lambda_i + k)^2} = p. \tag{17}$$

Με βάση, συνεπώς, τη σχέση (17) μπορεί να βρεθεί η τιμή του k.

Το πλεονέκτημα αυτής της μεθόδου είναι ότι η τιμή του k προσδιορίζεται από τη λύση της (17), αφού τα p και λ_i είναι γνωστά ως το πλήθος των ερμηνευτικών μεταβλητών και οι τιμές των χαρακτηριστικών ριζών της μήτρας συσχετίσεως $X'X$ αντίστοιχα. Ενώ στις μεθόδους της προηγούμενης παραγράφου ο προσδιορισμός του k είναι συνάρτηση των εκτιμήσεων OLS $\hat{\sigma}^2$ και $\hat{\beta}$ οι οποίες έχουν εκτιμηθεί από το αρχικό υπόδειγμά μας, δηλαδή από το υπόδειγμα με το πρόβλημα της πολυσυγγραμμικότητας.

Ακόμη ο Lee (1979) απέδειξε ότι με βάση το κριτήριο της ελαχιστοποιήσεως του MSE δεν υπερτερεί πάντοτε η RR της OLS. Χαρακτηριστικά σημειώνεται ότι στην «ιδανική περίπτωση» (ορθογώνιες μεταβλητές) συμβαίνει το αντίθετο. Ο Lee επιπλέον όρισε ως μέτρο του προβλήματος της πολυσυγγραμμικότητας τον ακόλουθο δείκτη

$$\mathbf{EFF} = \frac{\hat{\sigma}^2 \operatorname{tr}(X'X)^{-1} - \hat{\sigma}^2 \operatorname{tr}(VIF)}{(\text{Bias } (\hat{\beta}^*))^2}$$

$$= \frac{\hat{\sigma}^2 \sum_{i=1}^{p} \lambda i^{-1} - \sum_{i=1}^{p} \frac{\lambda_i}{(\lambda_i + k)^2}}{k^2 \sum_{i=1}^{p} \frac{\alpha_i^2}{(\lambda_i + k)}}. \tag{18}$$

Στη συνέχεια θα αναζητήσουμε μια μαθηματική έκφραση ανάμεσα στα MSE των OLS και RR και του δείκτη του Lee που δίνεται από την (18).

Πράγματι είναι

$$\text{MSE}(\hat{\beta}) - \text{MSE}(\hat{\beta}^*) = \sum_{i=1}^{p} \text{var}(\hat{\beta}^*) - \sum_{i=1}^{p} \text{var}(\hat{\beta}^*) - [\text{Bias } (\hat{\beta}^*)]^2 = \text{EFF}.$$

$$[\text{Bias } (\hat{\beta}^*)]^2 - [\text{Bias } (\hat{\beta}^*)]^2 = (\text{EFF} - 1) [\text{Bias } (\hat{\beta}^*)]^2.$$

Η τελευταία σχέση μας οδηγεί στο συμπέρασμα ότι αν EFF>1 τότε συνεπάγεται ότι ισχύει MSE $(\hat{\beta}^*) <$ MSE $(\hat{\beta})$, δηλαδή η σχέση (9), ενώ στην περίπτωση που ισχύει EFF<1 τότε MSE $(\hat{\beta}^*) >$ MSE $(\hat{\beta})$.

Ακόμη αν τεθεί $k = k_{max}$, για $EFF(k) = 1$, τότε η σχέση (9) δεν ισχύει για κάθε k αλλά για $0<k<k_{max}$. Επομένως κατά τον Lee η RR υπερτερεί της OLS με βάση το κριτήριο ελαχιστοποιήσεως του MSE μόνο για εκείνη την τιμή του k που ικανοποιεί τη σχέση $0<k<k_{max}$.

4. Εμπειρική αξιολόγηση των μεθόδων OLS, HKB, HSL, DSW, LW και VNC

Έστω το οικονομετρικό υπόδειγμα

$$C_t = \beta_0 + \beta_1 C_{t-1} + \beta_2 M_t + \beta_3 Y_t + u_t, \tag{19}$$

όπου C_t είναι η κατά κεφαλή κατανάλωση κατά χρονική περίοδο t, C_{t-1} η κατανάλωση της προηγούμενης περιόδου, M_t η κατά κεφαλή προσφορά χρήματος και Y_t το κατά κεφαλή εισόδημα. Τα στοιχεία που χρησιμοποιήθηκαν για την εκτίμηση του οικονομετρικού υποδείγματος (19) είναι από την Ελληνική Οικονομία για τη χρονική περίοδο 1953 – 1980 και αναφέρονται σε σταθερές τιμές του 1970.

Οι τιμές των δεικτών και κριτηρίων μετρήσεως της πολυσυγγραμμικότητας του πίνακα 1 για όλες τις μεθόδους προσδιορισμού του k δίνονται στον πίνακα 2, ενώ οι εκτιμήσεις των παραμέτρων β_1 και του δείκτη EFF δίνονται στον πίνακα 3.

Πίνακας 2
Τιμές των δεικτών και κριτηρίων μετρήσεως της πολυσυγγραμικότητας

	Μέθοδος	Τιμή k	M	max VIF	max R^2
1	OLS	.00000	212.26	104.517	.9904
2	HKB	.000308	188.14	92.449	.9891
3	HSL	.000157	199.42	98.086	.9892
4	DSW	.000309	188.06	92.406	.9891
5	LW	.000163	198.91	97.831	.9097
6	VNC	.09750	3.00	1.63	.3865

Πίνακας 3

Εκτιμήσεις των παραμέτρων του υποδείγματος

$$C_t = \beta_0 + \beta_1 C_{t-1} + \beta_2 M_t + \beta_3 Y_t + u_t$$

	Μέθοδος	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	EFF	R^2
1	OLS	.4402* (.0781)	−:0034 (.0365)	.4411* (.0551)	—	.9986
2	HKB	.3943* (.0722)	−.0013 (.0362)	.4444* (.0514)	12.5121	.9985
3	HSL	.3909* (.0741)	−.0019 (.0362)	.4470* (.0529)	24.177	.9986
4	DSW	.3943* (.0722)	−.0013 (.0362)	.4444* (.0514)	12.4655	.9985
5	LW	.3911* (.0743)	−.0019 (.362)	.4469* (.0528)	23.2265	.9986
6	VNC	.3234* (.0044)	.2421* (.0253)	.4434* (.0062)	0.2775	.9942

Οι αριθμοί εντός των παρενθέσεων είναι τα τυπικά σφάλματα και οι αστερίσκοι (*) δηλώνουν την στατιστική σημαντικότητα.

Από όλες τις τιμές των δεικτών M, maxVIF και maxR^2 υπάρχει ένδειξη για έντονο πρόβλημα πολυσυγγραμικότητας. Επίσης μία επιπλέον ένδειξη του προβλήματος είναι το αρνητικό αλγεβρικό πρόσημο της $\hat{\beta}_2$ πράγμα που δεν είναι συνεπές με την οικονομική θεωρία και ακόμη η $\hat{\beta}_2$ είναι στατιστικά ασήμαντη.

Οι μέθοδοι HKB, HSL, DSW, LW δεν επιτυγχάνουν θεραπεία του προβλήματος της πολυσυγγραμικότητας αφού οι τιμές των M, maxVIF και max R^2 παραμένουν αρκετά υψηλές και μακριά από τις «ιδανικές τιμές« (βλέπε πίνακα 1). Επίσης το πρόβλημα του πρόσημου της παραμέτρου β_2 εξακολουθεί να υφίσταται. Παρ' όλα αυτά όμως για όλες τις μεθόδους η τιμή του δείκτη EFF είναι μεγαλύτερη της μονάδας, δηλαδή σύμφωνα με το κριτήριο ελαχιστοποιήσεως του MSE οι μέθοδοι αυτές υπερτερούν όλες έναντι της μεθόδου OLS.

Αντίθετα η μέθοδος VNC δίνει στους δείκτες M, max VIF max R^2 σχεδόν τις «ιδανικές τιμές» (πίνακας 1), δηλαδή το πρόβλημα της πολυσυγγραμικότητας δεν υφίσταται πλέον. Όλες οι τιμές των β είναι θετικές, στατιστικά σημαντικές, σύμφωνες προς την οικονομική θεωρία και με μικρότερα τυπικά σφάλματα.

Η τιμή του k που επιτυγχάνεται με την VNC είναι υψηλότερη έναντι όλων των άλλων μεθόδων. Εδώ θα μπορούσε κανείς να πει ότι αν και μικρές τιμές k είναι επιθυμητές, με την έννοια ότι γίνεται μικρότερη μεταβολή στα αρχικά δεδομένα του υποδείγματος, εντούτοις οι τιμές αυτές δεν μειώνουν αρκετά το πρόβλημα της πολυσυγγραμμικότητας. Επίσης για την μέθοδο VNC η τιμή του δείκτη EFF είναι μικρότερη της μονάδας, δηλαδή σύμφωνα με το κριτήριο ελαχιστοποιήσεως του MSE στην περίπτωση αυτή υπερτερεί η OLS.

Τέλος γεννιέται το ερώτημα αν το πρόβλημα της πολυσυγγραμμικότητας παύει να υφίσταται όταν παραλείψουμε τη μεταβλητή M_t από το υπόδειγμα (19) της οποίας η εκτίμηση $\hat{\beta}_2$ είναι προβληματική από πλευράς προσήμου και στατιστικής σημαντικότητας. Στην περίπτωση αυτή το υπόδειγμα (19) γράφεται

$$C_t = \beta_0 + \beta_1 C_{t-1} + \beta_3 Y_t + u_t. \tag{20}$$

Οι τιμές των δεικτών μετρήσεως της πολυσυγγραμμικότητας, δίνονται στο πίνακα 4 ενώ οι εκτιμήσεις δίνονται στον πίνακα 5.

Πίνακας 4

Τιμές των δεικτών μετρήσεως της πολυσυγγραμμικότητας

	Μέθοδος	Τιμή k	M	max VIF	max R^2
1	OLS	.00000	207.77	103.88	.9904
2	HKB	.000197	191.79	95.89	.9947
3	HSL	.000104	199.08	99.54	.9949
4	DSW	.000198	110.43	95.21	.9903
5	LW	.000105	119.06	99.53	.9949
6	VNC	.051437	2.0005	1.0002	.0003

Με την παράλειψη της μεταβλητής M_t παρατηρούμε από τους πίνακες 2, 3, 4 και 5 τα εξής:

(α) Οι τιμές των k μεταβλήθηκαν και πιο ευαίσθητη παρουσιάζεται η μέθοδος VNC. (β) Οι δείκτες M, max VIF και max R^2 δεν μεταβλήθηκαν σημαντικά. (γ) Οι εκτιμήσεις της $\hat{\beta}_1$ παρουσιάζουν κάποια ευστάθεια σε σύγκριση με εκείνες που προκύπτουν από το υπόδειγμα που περιείχε την M_t, μεγαλύτερη όμως ευστάθεια παρουσιάζει η $\hat{\beta}_3$. (δ) Μόνο η μέθοδος VNC φαίνεται να θεραπεύει το πρόβλημα της πολυσυγγραμμικότητας και τέλος για τον δείκτη EFF ισχύουν όσα αναφέρθηκαν στον πίνακα 3.

Πίνακας 5
Εκτιμήσεις του υποδείγματος
$C_t = \beta_0 + \beta_t C_{t-1} + \beta_3 Y_t + u_t$

	Μέθοδος	$\hat{\beta}_1$	$\hat{\beta}_3$	EFF	R^2
1	OLS	.3868*	.4495*	—	.9986
		(.0750)	(.0543)		
2	HKB	.3916*	.4461*	18.6170	.9985
		(.0712)	(.0513)		
3	HSL	.3894*	.4477*	34.8611	.9986
		(.0735)	(0.522)		
4	DSW	.3913*	.4461*	18.5596	.9985
		(.0721)	(.0513)		
5	LW	.3894*	.4477*	34.7827	.9986
		(.0735)	(.0522)		
6	VNC	.3607*	.4862*	0.4376	.9977
		(.0052)	(.0074)		

Οι αριθμοί εντός των παρενθέσεων είναι τα τυπικά σφάλματα και οι αστερίσκοι (*) δηλώνουν την στατιστική σημαντικότητα.

5. Τεχνική MONTE-CARLO

Στην προηγούμενη παράγραφο έγινε εμπειρική αξιολόγηση των διαφόρων μεθόδων προσδιορισμού του k στη μέθοδο RR. Με την τεχνική Monte – Carlo θα γίνει η σύγκριση τόσο των μεθόδων προσδιορισμού του k όσο και των MSE των εκτιμητριών που επιτυγχάνονται μ' αυτές τις τιμές του k.

Για την περίπτωση αυτή επιλέγουμε την εξής προσέγγιση. Στα υποδείγματα (19) και (20) δίνονται οι εξής τιμές στους συντελεστές: $\beta_0 = 0$, $\beta_1 = 0.5$, $\beta_2 = 0.01$ και $\beta_3 = 0.4$, οι οποίες είναι κατά προσέγγιση οι τιμές που προκύπτουν με τη μέθοδο RR σύμφωνα με τη τεχνική των Hoerl and Kennard (1970). Επίσης είναι $u_t \sim N(0,\sigma^2)$ με $\sigma^2 = 0.0003$ η οποία υπολογίστηκε σύμωνα με τον Thisted (1976) και είναι σταθερή για όλες τις επαναλήψεις ώστε να έχουμε μικρό υπολογιστικό κόστος. Οι επαναλήψεις που έγιναν είναι 500 με νέα τυχαία σφάλματα για κάθε επανάληψη από την κατανομή N(0,.0003). Στη συνέχεια υπολογίστηκαν οι τιμές του k για όλες τις μεθόδους και το συνολικό μέσο σφάλμα τετραγώνου (TMSE). Σαν κριτήριο συγκρίσεως ορίστηκε ο λόγος

$$p_m = \frac{\text{TMSE (m)}}{\text{TMSE (OLS)}}, \quad m = \text{OLS, HKB, HSL, DSW, LW, VNC} \tag{21}$$

όπου

$$\mathrm{TMSE}(m) = \frac{1}{500} \sum_{i=0}^{3} \sum_{j=1}^{500} (\beta^m_{ij}(k) - \beta_i)^2 \qquad (22)$$

και $\beta^m_{ij}(k)$ = η εκτιμήτρια της i παραμέτρου η οποία επετεύχθει στη j επανάληψη με τη μέθοδο m.

Τα αποτελέσματα των 500 επαναλήψεων της τεχνικής Monte – Carlo για τα υποδείγματα (19) και (20) των τιμών k, του λόγου p_m και του δείκτη EFF δίνονται στον πίνακα 6.

Πίνακας 6

Τιμές των k, p_m και EFF

		$C_t = \beta_0 + \beta_1 C_{t-1} + \beta_2 M_t + \beta_3 Y_t + u_t$			$C_t = \beta_0 + \beta_1 C_{t-1} + \beta_3 Y_t + u_t$		
	Μέθ.	k	p_m	EFF	k	p_m	EFF
1	OLS	.0000 (.00000)	1.0000	—	.00000 (.00000)	1.000	—
2	HKB	.000245 (.000072)	0.9495	12.5121	.000160 (.000046)	0.9693	18.6170
3	HSL	.000127 (.000036)	0.9653	24.177	.000086 (.000024)	0.9779	34.8611
4	DSW	.000246 (.000073)	0.9490	12.4655	.000160 (.000046)	0.9651	18.5596
5	LW	.000132 (.000038)	0.9640	23.2256	.000086 (.000024)	0.9779	34.7827
6	VNC	.097500 (.0000)	9.7470	0.2775	.051473 (.00000)	3.5175	0.4376

Οι αριθμοί των παρενθέσεων είναι τα τυπικά σφάλματα και οι αστερίσκοι (*) δηλώνουν την στατιστική σημαντικότητα.

Από τον πίνακα 6 προκύπτει ότι η τιμή του k, η οποία προσδιορίστηκε από τις 500 επαναλήψεις της τεχνικής Monte – Carlo, με τη μέθοδο HKB είναι ακριβώς η ίδια με εκείνη της μεθόδου DSW. Επίσης οι τιμές των HLS και LW ελάχιστα διαφέρουν, η διαφορά τους είναι μικρότερη του .000005. Η μέθοδος DSW (ή HKB) δίνει μεγαλύτερες τιμές στο k απ' ότι η HSL (ή LW). Αντίθετα η μέθοδος VNC δίνει τη μεγαλύτερη τιμή στο k, μάλιστα είναι κατά 750 φορές μεγαλύτερη της τιμής του k της μεθόδου HSL. Από τη σύγκριση των τιμών του p_m προκύπτει ότι οι μέθοδοι HKB, HSL, DSW, και LW δίνουν τιμή μικρότερη της OLS και μάλιστα η DSW δίνει τη μικρότερη τιμή όλων. Δηλαδή με βάση το κριτήριο ελαχιστοποιήσεως του MSE όλες υπερέχουν έναντι της OLS και μάλιστα η DSW υπερέχει όλων. Αντίθετα η τιμή του p_m που επιτυγχάνεται με τη μέθοδο VNC είναι η μεγαλύτερη όλων, πράγμα που φανερώνει ότι η μέθοδος αυτή αν και δίνει «καλύτερα» αποτελέσματα από πλευράς εκτιμήσεως και θεραπείας του προβλήματος της πολυσυγγραμμικότητας (βλέπε πίνακες 2 και 3), εντούτοις με κριτήριο την ελαχιστοποίηση του MSE υστερεί έναντι όλων των άλλων μεθόδων. Συγκρίνοντας την τιμή του δείκτη p_m των δύο υποδειγμάτων προκύπτει ότι η προσθήκη της M_t βελτίωση τη μέθοδο RR σε σχέση με την OLS, το αντίθετο συμβαίνει με τη μέθοδο VNC. Τέλος με κριτήριο τον δείκτη EFF επιβεβαιώνεται ότι υπάρχει k για το οποίο ισχύει MSE (β*) >MSE (β).

6. Συμπεράσματα

Η μέθοδος RR που προτάθηκε από τους Hoerl και Kennard για τη θεραπεία του προβλήματος της πολυσυγγραμμικότητας μας δημιουργεί τα εξής προβλήματα: (α) ο προσδιορισμός της σταθεράς k η οποία υπεισέρχεται στην εκτιμήτρια Ridge, όπως προκύπτει από τη σχέση (7), εξαρτάται από τις «προβληματικές» εκτιμήσεις των β και σ^2 με την έννοια ότι προέκυψαν από το αρχικό υπόδειγμα που παρουσίαζε έντονο πρόβλημα πολυσυγγραμμικότητας και (β): το κριτήριο ελαχιστοποιήσεως του MSE δηλαδή MSE (β*)KM8E(β) δεν ισχύει για κάθε τιμή στη σταθερά k. Τούτο αποδείχθη στην περίπτωση ενός ορθογωνίου συστήματος («ιδανική περίπτωση»).

Κοινό σημείο των περισσοτέρων μεθόδων που έχουν προταθεί και μετά τις μεθόδους που χρησιμοποιούνται στη παρούσα εργασία (Casella (19985)) είναι η αποδοχή της RR και η προσπάθεια βελτιώσεως του μηχανισμού προσδιορισμού της σταθεράς k.

Η μέθοδος VNC που προτάθηκε από τον Lee φαίνεται να λύνει μερικώς τα παραπάνω προβλήματα. Ο προσδιορισμός, σύμφωνα με τη μέθοδο του Lee της k επιτυγχάνεται, όπως προκύπτει από την σχέση (17), χωρίς να χρησιμοποιούμε εκτιμήσεις των β και σ^2 που προκύπτουν από το αρχικό, παθολογικό υπόδειγμα. Η εισαγωγή του δείκτη EFF δίνει λύση στη χρήση του κριτηρίου ελαχιστοποιήσεως του MSE για $0<k<k_{max}$. Πρέπει όμως να τονιστεί ότι, με τη σειρά του ο δείκτης EFF εξαρτάται από τις εκτιμήσεις των β και σ^2 του αρχικού υποδείγματος. Έγινε επίσης φανερό ότι μόνο η μέθοδος VNC θεραπεύει ριζικά το πρόβλημα της πολυσυγγραμμικότητας.

Τέλος από την εφαρμογή της τεχνικής Monte - Carlo προέκυψαν τα εξής συμπεράσματα:

Οι τιμές της σταθεράς k των μεθόδων HKB, HSL, DSW, LW είναι μικρές και δεν διαφέρουν ουσιαστικά μεταξύ τους. Όλες οι μέθοδοι αυτές υπερτερούν της OLS με βάση το κριτήριο ελαχιστοποιήσεως του MSE και μάλιστα σύμφωνα με αυτό το κριτήριο η DSW υπερτερεί όλων. Εντούτοις όμως, σύμφωνα με τις τιμές των δεικτών μετρήσεως της πολυσυγγραμμικότητας, το πρόβλημα της πολυσυγγραμμικότητας εξακολουθεί να υπάρχει και μετά την εισαγωγή της σταθεράς κ.

Αντίθετα αποτελέσματα επιτυγχάνονται με τη μέθοδο VNC. Η τιμή του k είναι ουσιαστικά μεγαλύτερη όλων των άλλων τιμών. Με βάση το κριτήριο ελαχιστοποιήσεως του MSE υστερεί έναντι όλων των άλλων μεθόδων ακόμη και της OLS. Οι τιμές όμως των δεικτών μετρήσεως της πολυσυγγραμμικότητας, εγγίζουν τις «ιδανικές τιμές», δηλαδή θεραπεύεται το πρόβλημα της πολυσυγγραμμικότητας.

Κλείνοντας θα μπορούσαμε να παρατηρήσουμε ότι το κλειδί στη μέθοδο RR παραμένει ο τρόπος προσδιορισμού του k και του πεδίου τιμών του.

ΒΙΒΛΙΟΓΡΑΦΙΑ

Casella, G. (1985), "Condition numbers and minimax ridge regression estimators.journal *f the American Statistical Association, 753 - 758.*

Dempster A. P., Schaltzoff M. and *Wermuth, N.* (1977), "A Simulation study of alternatives to ordinary least squares". Zournal of the American Statistical Association (with discussion), 72, 77 - 106.

Hocking, R. R., Speed, F.M. and *Lynn, M.J.* (1976), "A class of biased estimators in linear regression. Technometrics, 18, 425-438.

HoerJ, *A.E.* and *Kennard R. W.* (1970a), "Ridge regression: Biased estimation of nonorthogonal problems", Technometrics, 12, 55-67.

Hoerl, A. E. and *Kennard, R.W.* (1970b), "Ridge regression: Application to nonorthogonal problems". Technometrics, 12, 69-82.

Hoerl, A.E. and Kennard, R.W and *Baldwin, K.F.* (1975), "Ridge regression: Some simulations" Communications in Statistics, 4(2), 105- 125.

Kmenta, J. (1971), Elements of Econometrics, MacMillan, N. York - London.

Lawless, J.F. and *Wang, P.* (1976), "A simulation Study of ridge and other regression estimators". Communications in Statistics-Theory and Methods, A5(4), 307-323.

Lawless, J. F. (1978), "Ridge and related estimation procedures: theory and practice". Communications in Statistics-Theory and Methods, A7(2), 139-164.

Lee, W.F. (1979), "Model estimation using ridge regression with the variance normalization criterion". Interim report No2. The education and inequality in Canada project, university of Newfoundland.

Mahayan, V., Jain, A.K. and *Bergier, M.* (1977), "Parameter estimation in marketing models in the presence of multicollinearity. An application of ridge regression". Journal of Marketing Research, 14, 586-591.

Marquardt, D.W. and *Snee, R.D.* (1975), 'Ridge regression in practice". The American Statistician, 29, 3-19.

Thisted, R.A. (1976), "Ridge regression. Minimax Estimation and Empirical Bayes Methods", technical Report 87, Division of Biostatistics, Stanford University.

Thisted, R.A. (1980), "Comment on Smith and Campbell (1980)". Journal of the American Statistical Association, 75, 91-86.