To appear in the *Journal of Applied Statistics*
Vol. 00, No. 00, Month 20XX, 1–22

Collinearity: Revisiting the Variance Inflation Factor in Ridge Regression

C.B. García[a] and J. García[b] and M.D.M. López[a*] and R. Salmerón[a]

[a] *Department of Quantitative Methods for Economic and Business, University of Granada, Spain*; [b] *Department of Economic and Business, University of Almería, Spain*

(*v4.0 released November 2013*)

Ridge regression has been widely applied to estimate under collinearity by defining a class of estimators that are dependent on the parameter k. The Variance Inflator Factor (VIF) is applied to detect the presence of collinearity and also as an objective method to obtain the value of k in ridge regression. Contrarily to the definition of the variance inflator factor, the expressions traditionally applied in ridge regression do not necessarily lead to values of VIFs equal to or greater than 1. This work presents an alternative expression to calculate the variance inflator factor in ridge regression that satisfies the afore mentioned condition, and also presents other interesting properties.

Keywords: Collinearity; Ridge Regression; Variance Inflation Factor; Linear Regression; Covariance Matrix.

1. Introduction

The collinearity problem deals with the existence of linear relationships between two or more independent variables in the following linear model with n observations and p independent variables:

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{u}, \tag{1}$$

where it is supposed that it is correctly specified, that is the model has mean zero **(there is no y-intercept parameter β_0)**, homoscedasticity and uncorrelated error. If the collinearity is perfect, $\text{rg}(\mathbf{X}) < p$, **there is no unique solution to estimate** (1), while if it is approximate the estimation will be unstable.
Some of the possible solutions to the problem focus on the sample (to improve the sample, to design information extracting maximum observed variables or to increase the sample size). Another common solution is simply to dispense with the variable that produces collinearity. In this case it is possible to remove the relevant variables which can lead to heteroscedasticity and autocorrelation, etc. Thus, although this solution seems to be the most widespread, it may well not be the most appropriate.
Alternatively, some authors suggest treating the problem of collinearity mechanically, by proposing a technique known as ridge regression [16, 17], whose estimators are obtained as:

*Corresponding author. Email: mariadelmarlopez@ugr.es

$$\hat{\boldsymbol{\beta}}_R(k) = \left(\mathbf{X}'\mathbf{X} + k\mathbf{I}\right)^{-1} \mathbf{X}'\mathbf{Y}, \quad k \geq 0. \tag{2}$$

This technique has been widely studied and some authors have even tried to improve it with numerous proposals. On the other hand, different authors have presented some alternative techniques to ridge regression (e.g., the surrogate ridge model Jensen and Ramirez [21] recently cited by Hadi [12], the raise method García et al. [9] and the nested regression Feng-Jeng [6]).
A widely applied measure to analyze the problem of collinearity is the Variance Inflator Factor (VIF) which is defined in an standardized model as:

$$\text{VIF}\left(\hat{\beta}_i, \hat{\beta}_i^S\right) = \frac{\text{var}\left(\hat{\beta}_i\right)}{\text{var}\left(\hat{\beta}_i^S\right)}, \quad i = 1, \ldots, p. \tag{3}$$

where $\hat{\boldsymbol{\beta}}$ is an estimator of $\boldsymbol{\beta}$ and $\hat{\boldsymbol{\beta}}^S$ is the corresponding estimator under a standard model **(no collinearity)**. By considering:

$$\text{var}\left(\widehat{\beta}_i\right) = \frac{\sigma^2}{\sum\limits_{j=1}^{n} \left(X_{ji} - \overline{\mathbf{X}}_i\right)^2} \frac{1}{1 - R_i^2}, \quad i = 1, \ldots, p. \tag{4}$$

If the exogenous variables are orthogonal to the variable $\mathbf{X}_i$ then $R_i^2 = 0$ and the expression of the variance of the standard model estimator will be given by:

$$\text{var}\left(\widehat{\beta}_i^S\right) = \frac{\sigma^2}{\sum\limits_{j=1}^{n} \left(X_{ji} - \overline{\mathbf{X}}_i\right)^2}, \quad i = 1, \ldots, p. \tag{5}$$

Thus, it is possible to obtain the generally accepted definition of VIF due to Theil [38]:

$$\text{VIF}_i = \frac{1}{1 - R_i^2}. \tag{6}$$

Note that the second factor of expression (4) becomes the VIF and measures the impact of collinearity of the variable X_i, $i = 1, \ldots, p$, with the rest of variables on the square of the radius of the confidence interval.
According to Marquardt [29], the elements of the main diagonal of the matrix $(\mathbf{X}'\mathbf{X} + k\mathbf{I})^{-1} \mathbf{X}'\mathbf{X} (\mathbf{X}'\mathbf{X} + k\mathbf{I})^{-1}$ will be the variance inflation factors. Marquardt [29] and McDonald [32] recommended the use of the variance inflation factor as a procedure to select the value for k, by proposing the following expressions, respectively, for two independent variables in a standardized model:

$$\text{VIF}_M(k) = \frac{(1+k)^2 - 2(1+k)\rho^2 + \rho^2}{\left[(1+k)^2 - \rho^2\right]^2}, \tag{7}$$

and

$$\mathrm{VIF}_{McD}(k) = \frac{\lambda_1 (\lambda_1 + k)^{-2} + \lambda_2 (\lambda_2 + k)^{-2}}{2}, \tag{8}$$

where $\lambda_1 = 1+\rho$ and $\lambda_2 = 1-\rho$ are the latent roots of $\mathbf{X}'\mathbf{X}$. Note that after substitution, both expressions are equal, that is $\mathrm{VIF}_M(k) = \mathrm{VIF}_{McD}(k)$. Thus, in what follows we will refer only to the VIF proposed by Marquardt [29] as $\mathrm{VIF}_M(k)$.
On the other hand, O'Brien [36] states that the VIF *is a measure of the* ith *independent variables collinearity with the other independent variables in the analysis and is connected directly to the variance of the regression coefficient associated with this independent variable*. Thus, the unbiased estimate of the variance of the ith regression coefficient is given by [7, 11, 36]:

$$\widehat{\sigma}^2 \left(\hat{\boldsymbol{\beta}}_i\right) = \frac{\frac{(1-R_Y^2)\sum_{j=1}^n (Y_j - \bar{\mathbf{Y}})^2}{n-p-1}}{(1-R_i^2)\sum_{j=1}^n x_j^2}, \quad i = 1, \ldots, p. \tag{9}$$

This expression shows that the variance is influenced by different elements (the number of observations, the determination coefficient, the sample dispersion, the VIF, etc.). This fact implies that taking the elements of the main diagonal of the matrix $(\mathbf{X}'\mathbf{X} + k\mathbf{I})^{-1} \mathbf{X}'\mathbf{X} (\mathbf{X}'\mathbf{X} + k\mathbf{I})^{-1}$ as the VIF seems not to be appropriate as the variance is not only composed by the VIF but also by other elements.
Some recent works have analyzed and proposed some improvements and alternatives for the VIF as a measure of the existence of collinearity. For example, Dias and Castro [4] propose new indicators to this problem in the multiple linear regression model. Dias and Castro [5] show that the real impact on variance can be overestimated by the traditional VIF when the explanatory variables contain no redundant information about the dependent variable and propose corrected version of this collinearity indicator. Jensen and Ramirez [24] re-examine the VIFs for models with intercept, with and without centering regressors, and they suggest other VIFs in rotation of the conventional VIF. Our contribution is also focused on the variance inflation factor but particularly in its application within the ridge regression.
In this work we show that the expressions proposed by Marquardt [29] and McDonald [32] for calculating VIF **for** $\mathbf{p = 2}$ are not true VIFs although they have been applied as such in ridge regression. Thus, an alternative expression for the VIF is proposed in this paper. In order not to get into the controversy of the standardization we will work with standardized data. The structure of the paper is as follows: Section 2 includes some features of ridge estimatios and the used notation; in Section 3 we explain why the different expressions commonly applied to calculate the VIF in ridge regression are not correct justifying the presentation of an alternative expression; Section 4 shows the ridge regression VIF, $\mathrm{VIF}_R(k,n)$, as the main contribution of this paper. We show how to calculate it from the matrix that generates the ridge estimators and explain why $\mathrm{VIF}_R(k,n)$ is the true ridge regression VIF; Section 5 analyzes the asymptotic limit of $\mathrm{VIF}_R(k,n)$ showing that when n tends to ∞ the $\mathrm{VIF}_R(k,n)$ coincides with the surrogate ridge regression VIF presented by Jensen and Ramirez [21] and denoted by VIF_{Sk}; Section 6 presents a new expression for the variance deflator factor in the ridge regression. In Section 7 we present a numerical example to illustrate the contributions of this paper and, finally, Section 8 summarizes the conclusions and recommends the application of the $\mathrm{VIF}_R(k,n)$ or the VIF_{Sk} [21] to detect collinearity in ridge regression instead of the $\mathrm{VIF}_M(k)$ [29].

2. Ridge estimator properties. Notation

For convenience, it is assumed that the variables in $\mathbf{X}$ are standardized so that $\mathbf{X}'\mathbf{X}$ has the form of a correlation matrix. The estimators are:

- $\hat{\boldsymbol{\beta}}_L = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}$ in OLS regression.
- $\hat{\boldsymbol{\beta}}_R(k) = \left\{(\mathbf{X}'\mathbf{X}+k\mathbf{I})^{-1}\mathbf{X}'\mathbf{Y}, \quad k \geq 0\right\}$ in ridge regression (see Hoerl [14, 15]).

Some features of the ridge estimator are its coincidence with the Ordinary Least Square (OLS) estimator for $k = 0$ and the decrease of the determination coefficient as k increases, that is, R^2 is a decreasing function of k [32]. It is also a biased estimator of $\boldsymbol{\beta}$ when $k > 0$ and its covariance matrix is:

$$\operatorname{var}\left(\widehat{\boldsymbol{\beta}}_R(k)\right) = \sigma^2\left(\mathbf{X}'\mathbf{X}+k\mathbf{I}\right)^{-1}\mathbf{X}'\mathbf{X}\left(\mathbf{X}'\mathbf{X}+k\mathbf{I}\right)^{-1}. \tag{10}$$

Remember that Marquardt [29] takes the elements of the main diagonal of the matrix $(\mathbf{X}'\mathbf{X}+k\mathbf{I})^{-1}\mathbf{X}'\mathbf{X}(\mathbf{X}'\mathbf{X}+k\mathbf{I})^{-1}$ as the variance inflation factors. The different VIFs expressions will be denoted as follows for $k \geq 0$:

- $\text{VIF}_M(k)$ is the VIF presented by Marquardt [29] in the ridge regression.
- VIF_{Sk} is the VIF of the surrogate ridge model presented by Jensen and Ramirez [21].
- $\text{VIF}_R(k,n)$ is the VIF obtained from the matrix that generates the ridge estimators from the OLS regression.

Note that as the determination coefficient R_i^2 varies from 0 to 1, the VIF always has to be greater than 1. Then, expression (7) seems not to be appropriate because the calculation of the VIF in ridge regression can present values lower than 1. However, this expression has been repeatedly applied since Marquardt [29] proposed it. Even Fox and Monette [8] followed this same idea to obtain the generalized VIF. For this reason it is possible to find many examples in the existing literature where the VIF presents some values lower than 1. Since a complete enumeration would be a tedious question and without relevance, Table 1 serves as an illustration of the presence of VIF lower than 1 in the literature. However, this does not mean that the VIFs are monotone decreasing as a function of the ridge parameter k.

3. Analysis of the variance inflator factor expressions

By taking the following standardized model:

$$y_j = \beta_1 x_{1j} + \beta_2 x_{2j} + v_j, \quad j = 1, \ldots, n, \tag{11}$$

where the following conditions are assumed: $\sum_{j=1}^n x_{1j} = 0$, $\sum_{j=1}^n x_{1j}^2 = 1$, $\sum_{j=1}^n x_{2j} = 0$, $\sum_{j=1}^n x_{2j}^2 = 1$, and $\sum_{j=1}^n x_{1j}x_{2j} = \rho$. Then:

$$\mathbf{X}'\mathbf{X} = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}, \tag{12}$$

and the covariance matrix will be:

$$\text{var}\left(\hat{\boldsymbol{\beta}}_L\right) = \sigma^2 \left(\mathbf{X}'\mathbf{X}\right)^{-1} = \sigma^2 \begin{pmatrix} \frac{1}{1-\rho^2} & \frac{-\rho}{1-\rho^2} \\ \frac{-\rho}{1-\rho^2} & \frac{1}{1-\rho^2} \end{pmatrix}, \quad (13)$$

where each value of the diagonal presents the variance of the corresponding parameter estimator. Thus, the variance inflator factor could be defined, when $p = 2$, as the corresponding element of the main diagonal of the matrix $(\mathbf{X}'\mathbf{X})^{-1}$. That is to say

$$\text{VIF} = \frac{1}{1-\rho^2}. \quad (14)$$

If we use this last definition of VIF in the ridge estimator with expression (2) and with the matrix $\left(\mathbf{X}'\mathbf{X} + k\mathbf{I}\right)^{-1} \mathbf{X}'\mathbf{X} \left(\mathbf{X}'\mathbf{X} + k\mathbf{I}\right)^{-1}$, then we obtain the expression (7) for the $\text{VIF}_M(k)$ [29]. Since $-1 \leq \rho \leq 1$, we can see that expression (14) will be always equal to or greater than 1, while expression (7) can take values lower than 1.
It is also evident that the VIF should increases as the correlation coefficient ρ increases. However, the $\text{VIF}_M(k)$ does not verify this condition as shown in Figure 1. Note that the $\text{VIF}_M(k)$ decreases for values of ρ higher than 0.9 (when collinearity is serious), even taking values less than one. Then, we can state that the $\text{VIF}_M(k)$ begins to decrease just when collinearity problems appear, which could be misleading. This fact contradicts the universally accepted VIF definition presented in (6).
When we are working with the Ordinary Least Square (OLS) estimator for $p = 2$, the expression (13) is verified by the elements in the main diagonal of $(\mathbf{X}'\mathbf{X})^{-1}$ and, in consequence, the VIF of every variable is the corresponding diagonal element of this matrix. However, if we apply this same idea, but working with the ridge estimator, we will obtain a definition of the VIF that does not coincide with its natural definition, see expression (6). This fact explains why the expression of the $\text{VIF}_M(k)$ [29, 32] is not a true VIF and justifies the presentation of an alternative VIF expression in the following section.
To finish this section we note that when $p > 2$ the main diagonal of the matrix $(\mathbf{X}'\mathbf{X})^{-1}$ will content in its mean diagonal the usual VIFs expression (6). Thus, for example when $p = 3$ it is verified

$$\text{VIF}_i = \frac{1}{1 - \frac{\rho_{ij}^2 + \rho_{ik}^2 - 2\rho_{ij}\rho_{ik}\rho_{jk}}{1-\rho_{jk}^2}} = \frac{1-\rho_{jk}^2}{1-\rho_{ij}^2-\rho_{ik}^2-\rho_{jk}^2+2\rho_{ij}\rho_{ik}\rho_{jk}}. \quad (15)$$

where ρ_{ij} is the correlation between $\mathbf{X}_i$ and $\mathbf{X}_j$. However, to take the VIFs as the main diagonal of the matrix $(\mathbf{X}'\mathbf{X}+k\mathbf{I})^{-1}\mathbf{X}'\mathbf{X}(\mathbf{X}'\mathbf{X}+k\mathbf{I})^{-1}$ will continue to be incorrect for the same reasons that when $p = 2$. Thus, although the extension to $p > 2$ has some similarities with the case $p = 2$, it exhibits important peculiarities that overcome the goal of this paper.

4. The ridge regression VIF: $\text{VIF}_R(k, n)$

As indicated before, the VIFs presented by Marquardt [29] and McDonald [32] are incorrectly calculated in the matrix $\left(\mathbf{X}'\mathbf{X} + k\mathbf{I}\right)^{-1} \mathbf{X}'\mathbf{X} \left(\mathbf{X}'\mathbf{X} + k\mathbf{I}\right)^{-1}$ leading to a wrong definition of VIF. Now, we will present an alternative expression of the VIF obtained from the matrix $\mathbf{X}_A$ calculated to obtain the ridge estimator by using OLS regression.

Marquardt [29] and, more explicitly, Zhang and Ibrahim [41] pointed out that the ridge estimator can be calculated by OLS regression from the matrix $\mathbf{X}_A$ as:

$$\widehat{\boldsymbol{\beta}}_R(k) = \left(\mathbf{X}'\mathbf{X} + k\mathbf{I}\right)^{-1}\mathbf{X}'\mathbf{Y} = \left(\mathbf{X}'_A\mathbf{X}_A\right)^{-1}\mathbf{X}'_A\mathbf{Y}_A, \quad k \geq 0, \tag{16}$$

where $\mathbf{X}_A = \begin{pmatrix} \mathbf{X} \\ \sqrt{k}\mathbf{I} \end{pmatrix}$ and $\mathbf{Y}_A = \begin{pmatrix} \mathbf{Y} \\ \mathbf{0} \end{pmatrix}$ being $\mathbf{I}$ the identity matrix and $\mathbf{0}$ the null vector both of order p. Evidently, $\left(\mathbf{X}'_A\mathbf{X}_A\right)^{-1} = \left(\mathbf{X}'\mathbf{X} + k\mathbf{I}\right)^{-1}$, but what is really relevant is that we now know the matrix $\mathbf{X}_A$ that has generated the matrix $\left(\mathbf{X}'\mathbf{X} + k\mathbf{I}\right)^{-1}$. Hence, we can calculate the determination coefficient between the independent variables and the VIF from its general definition.
By developing the matrix $\mathbf{X}_A$ with $p = 2$ we have:

$$\mathbf{X}_A = \begin{pmatrix} x_{11} & x_{21} \\ x_{12} & x_{22} \\ \vdots & \vdots \\ x_{1n} & x_{2n} \\ \sqrt{k} & 0 \\ 0 & \sqrt{k} \end{pmatrix} = \begin{pmatrix} z_{11} & z_{21} \\ z_{12} & z_{22} \\ \vdots & \vdots \\ z_{1n+2} & z_{2n+2} \end{pmatrix}, \tag{17}$$

from which we can estimate the model $z_{1j} = \beta_1 + \beta_2 z_{2j} + w_j$, with $j = 1, \ldots, n+2$ (without standardized variables) where $\sum_{j=1}^{n+2} z_{1j} = \sqrt{k}$ and $\sum_{j=1}^{n+2} z_{2j} = \sqrt{k}$, and obtain the coefficient of determination.
Then, the estimator of the parameter is obtained as:

$$\hat{\boldsymbol{\beta}} = \frac{1}{(n+2)(1+k) - k} \begin{pmatrix} \sqrt{k} + k\sqrt{k} - \rho\sqrt{k} \\ -k + (n+2)\rho \end{pmatrix}. \tag{18}$$

To calculate the determination coefficient we will calculate the Total Sum of Squares (TSS) and the Explained Sum of Squares (ESS):

$$\text{TSS} = 1 + k - (n+2)\left(\frac{\sqrt{k}}{n+2}\right)^2 = \frac{(1+k)(n+2) - k}{n+2}, \tag{19}$$

and

$$\text{ESS} = \frac{\left[(n+2)\rho - k\right]^2}{(n+2)\left[(n+2)(1+k) - k\right]}. \tag{20}$$

Then, the determination coefficient will be obtained as:

$$R_i^2 = \frac{\text{ESS}}{\text{TSS}} = \frac{\left[(n+2)\rho - k\right]^2}{\left[(n+2)(1+k) - k\right]^2}, \quad i = 1, 2. \tag{21}$$

The determination coefficient will be 0 when $k = (n+2)\rho$ and then the VIF will be 1. When $k = (n+2)\rho$ the variables are orthogonal and, from this value, the determination

coefficient will increase until a limit value equal to $\frac{1}{(n+1)^2}$ as k tends to infinite. See Figure 2.
And the VIF for the ridge regression will be obtained as:

$$\text{VIF}_R(k,n) = \frac{1}{1-R_i^2} = \frac{[(n+2)(1+k)-k]^2}{(n+2)^2\,[(1+k)^2-\rho^2] - 2(n+2)k(1+k-\rho)}. \qquad (22)$$

Note that expression (22) depends not only on the parameter k but also on the number of observations, n, as pointed out by O'Brien [36].

5. The surrogate ridge regression VIF, VIF_{Sk}, and its asymptotic equivalence with $\text{VIF}_R(k,n)$

When we use the Ordinary Least Square (OLS) estimation in the model (1), the VIF is obtained by using the expression (6) and the matrix $\mathbf{X}$ of the model (1). On the other hand, if we are working with the ridge regression the determination coefficient to obtain the VIF from its traditional definition could be calculated of different way from the matrix $\mathbf{X}_k$ presented by Jensen and Ramirez [21, 22] which verifies $\mathbf{X}_k'\mathbf{X}_k = \mathbf{X}'\mathbf{X} + k\mathbf{I}$, see Table 2. Then, it is possible to show that for $p = 2$:

$$\mathbf{M} = \mathbf{X}_k'\mathbf{X}_k = \begin{pmatrix} m_{11} \dots m_{1n} \\ m_{21} \dots m_{2n} \end{pmatrix} \begin{pmatrix} m_{11} & m_{21} \\ \vdots & \vdots \\ m_{1n} & m_{2n} \end{pmatrix} = \mathbf{X}'\mathbf{X} + k\mathbf{I} = \begin{pmatrix} 1+k & \rho \\ \rho & 1+k \end{pmatrix}. \qquad (23)$$

In this way, given the moment matrix $\mathbf{M}$, it is obtained the VIF of the surrogate ridge regression [21] denoted by VIF_{Sk}, for the variable i with $i = 1, 2$:

$$\text{VIF}_{Sk} = \frac{\left(\mathbf{M}^{-1}\right)_{ii}}{\frac{1}{\mathbf{M}_{ii}}} = \left(\mathbf{M}^{-1}\right)_{ii}\mathbf{M}_{ii} = \frac{1+k}{(1+k)^2-\rho^2}(1+k). \qquad (24)$$

Note that this expression does not coincide with the expression (7) proposed in Marquardt [29] and McDonald [32]. We can prove that expression (24) is always greater than 1, an increasing function respect to ρ^2 and it coincides with the VIF in the OLS estimator when k is equal to zero. The matrices $\mathbf{X}_k$ and $\mathbf{X}_A$ have different dimensions. For this reason, the VIF_{Sk} and $\text{VIF}_R(k,n)$ will not be equal. Then, now we will analyze the asymptotic equivalence between them.
The asymptotic VIF is defined as the limit of the $\text{VIF}_R(k,n)$ when n tends to infinity. In the present case of two independent variables, it could be shown that:

$$\text{VIF}_R(k,\infty) = \lim_{n\to\infty} \text{VIF}_R(k,n) = \frac{(1+k)^2}{(1+k)^2-\rho^2}. \qquad (25)$$

When n tends to infinity the asymptotic ridge VIF coincides with the VIF_{Sk} presented in expression (24). Indeed, it can be proved that the $\text{VIF}_R(k,n)$ practically coincides with the VIF_{Sk} for values of n higher than 30 (see Figure 3).
On the other hand, as k tends to infinity, the limit of $\text{VIF}_R(k,n)$ is equal to:

$$\text{VIF}_R(\infty, n) = \lim_{k \to \infty} \text{VIF}_R(k, n) = \frac{(n+1)^2}{n(n+2)}. \tag{26}$$

Thus, the $\text{VIF}_R(k, n)$ will be always greater than 1 and tends to 1 as n and k tend to infinity. Although the scientific literature suggests that the parameter k varies in the interval $[0, \infty)$, the ridge regression solutions do not improve the OLS solutions in all cases. The MSE curve for ridge regression crosses the OLS curve from below as k increases. For this reason, for suitable values of k and under the conditions cited by Jensen and Ramirez [23], the OLS solutions eventually dominate in MSE the ridge regression solution.

In order to analyze the difference between the $\text{VIF}_R(k, n)$ and the VIF_{Sk}, it is calculated as:

$$\text{VIF}_{Sk} - \text{VIF}_R(k, n) = \frac{k\left(-k^2 - k + 4\rho + 3k\rho + 2n\rho + 2nk\rho\right)}{(2+n)\left[(1+k)^2 - \rho^2\right](2 + n + kn + 2\rho + n\rho)}. \tag{27}$$

It can be empirically demonstrated that it is possible to use the VIF_{Sk} instead of the $\text{VIF}_R(k, n)$ for values of n higher than 30 and so the difference will be inappreciable. In addition, both expressions present values higher than the traditionally applied VIF expressions [29] and [32].

It is possible to demonstrate that the VIF_{Sk} is equal to the $\text{VIF}_R(k, n)$ when:

$$n = \frac{k^2 + k - 3k\rho - 4\rho}{2(1+k)\rho}. \tag{28}$$

This expression takes negative values when $|\rho| > 0.5$ and k belongs to $[0, 1]$. From this evidence, we can conclude that the VIF_{Sk} and the $\text{VIF}_R(k, n)$ do not intersect for any value of $n > 0$ and $|\rho| > 0.5$.

Now, we will study from an analytical point of view the expression (27) **by supposing that n is a real number higher or equal than zero**. When $n = 0$ the expression (27) is:

$$\text{VIF}_R(k, 0) - \text{VIF}_{Sk} = \frac{k^2(1+k) - 3\rho k(1+k) - \rho k}{\left[(1+k)^2 - \rho^2\right] 4(1+\rho)}. \tag{29}$$

Note that when ρ is negative, the numerator will be positive and the $\text{VIF}_R(k, 0)$ is higher than the VIF_{Sk}. On the contrary, when ρ is positive the numerator will be negative and the $\text{VIF}_R(k, 0)$ is less than the VIF_{Sk} since ρ should be higher than $\frac{k^3+k^2}{3k^2+4k}$ to verify $k^2(1+k) - 3\rho k(1+k) - \rho k > 0$. Figure 4 shows, for $k = 0.3$, the $\text{VIF}_R(0.3, \infty)$, the $\text{VIF}_R(0.3, n)$ for $\rho = 0.95$, the $\text{VIF}_R(0.3, n)$ for $\rho = -0.95$ and the $\text{VIF}_M(0.3)$. Note that the $\text{VIF}_M(0.3)$ presents a constant value less than 1 while the $\text{VIF}_R(0.3, \infty)$ is also constant but higher than 1. The $\text{VIF}_R(0.3, n)$ with $\rho = -0.95$ is always higher than the $\text{VIF}_R(0.3, \infty)$ while the $\text{VIF}_R(k, n)$ with $\rho = 0.95$ is always less. Both of them ($\text{VIF}_R(0.3, n)$ with $\rho = -0.95$ and $\text{VIF}_R(0.3, n)$ with $\rho = 0.95$) converge to the $\text{VIF}_R(k, \infty)$ with $k = 0.3$.

It can be shown that

$$\frac{\partial \text{VIF}_R(k,n)}{\partial n} = \frac{[(n+2)(1+k)-k]\,2k(1+k-\rho)\,[(n+2)\rho-k]}{[(n+2)^2\,[(1+k)^2-\rho^2]-2(n+2)k(1+k-\rho)]^2}. \quad (30)$$

Then, it is possible to conclude that for values of ρ between -1 and -0.9 the partial derivative in (30) will be negative, the $\text{VIF}_{\mathbf{R}}(\mathbf{k},\mathbf{n})$ will be a decreasing function and the difference between the $\text{VIF}_R(k,n)$ and the VIF_{Sk} (27) will be positive. The opposite case will be presented for values of ρ between 0.9 and 1. Note that the sign of the expression (30) depends on the factor $(n+2)\rho-k$. See Table 3.

6. The ridge variance deflation factor: $\mathbf{VDF}_{R}(k)$

O'Brien [36] established the concept of variance deflator factor in the context of OLS regression from expression (9) as:

$$\text{VDF} = 1 - R^2_{\mathbf{Y}}, \quad (31)$$

where $R^2_{\mathbf{Y}}$ is the determination coefficient of the regression. McDonald [32] provided an explicit expression of $R^2_{\mathbf{Y}}$ for the case $p=2$, $k=0$, $\rho>0$ and standardized variables given by:

$$R^2_{\mathbf{Y}} = \frac{(\gamma_1-\gamma_2)^2 + 2(1-\rho)\gamma_1\gamma_2}{1-\rho^2}, \quad (32)$$

where γ_1 and γ_2 are the correlation coefficients of the response variable with each of the exogenous variables $\mathbf{X}_1$ and $\mathbf{X}_2$, respectively.
The extension of the VDF established by O'Brien [36] to the ridge regression will make no sense since $R^2_{\mathbf{Y}}$ is decreasing in k [32] and then expression (31) will be increasing in k contrarily to the expected result.
Now, a new expression for the VDF will be defined in the context of the ridge regression, called the ridge variance deflator factor and denoted by $\text{VDF}_R(k)$. This definition will depend on k contrarily to expression (31).
For this purpose, we start from the equivalence between the $\text{VIF}_R(k,\infty)$ and the VIF_{Sk} for $n>30$. Then, we could consider the VIF_{Sk} as a measure of the dependence between the exogenous variables. From expressions (10) and (7) we can obtain:

$$\hat{\sigma}^2(\hat{\beta}_i) = \hat{\sigma}^2 \frac{(1+k)^2 - 2(1+k)\rho^2 + \rho^2}{[(1+k)^2-\rho^2]^2}. \quad (33)$$

If we multiply and divide this expression by the VIF_{Sk}, we will obtain:

$$\hat{\sigma}^2(\hat{\beta}_i) = \hat{\sigma}^2 \frac{(1+k)^2 - 2(1+k)\rho^2 + \rho^2}{(1+k)^2\,[(1+k)^2-\rho^2]} \text{VIF}_{Sk} = \hat{\sigma}^2 \cdot \text{VDF}_R(k) \cdot \text{VIF}_{Sk}. \quad (34)$$

Considering the expression (34), we define the ridge variance deflator factor as:

$$\text{VDF}_R(k) = \begin{cases} \frac{(1+k)^2-2(1+k)\rho^2+\rho^2}{(1+k)^2[(1+k)^2-\rho^2]} & \text{, for } k > 0 \\ 1 & \text{, for } k = 0 \end{cases}. \tag{35}$$

It can be observed that expressions (31) and (35) are different even when $k = 0$.
The $\text{VDF}_R(k)$ will be always less than the VIF_{Sk}, since it can be proved that:

$$\frac{\text{VDF}_R(k)}{\text{VIF}_{Sk}} = \frac{k^2 + 2k(1-\rho^2) + (1-\rho^2)}{(1+k)^4} < 1, \tag{36}$$

when $-1 < \rho < 1$ and $k > 0$. Consequently, it is evident that this factor acts as a deflator of the variance. Furthermore, it could be proved that expression (36) is a decreasing function of k and ρ^2.

7. Empirical application

To illustrate the contribution of this paper, we will reproduce the example previously presented in [32, 33]. In this example the total mortality rate, Y, is related to the nitrogen oxide pollution potential, X_1, and the hydrocarbon pollution potential, X_2, for 60 cities. From this information, we can obtain the value of $n = 60$ and $\rho = 0.984$.
The focus of this article is not to discuss the selection of the ridge parameter, k, but to see when the problem of collinearity has been solved. Then, we will use Table 4 to directly present the values of the VIF_{Sk}, the $\text{VIF}_R(k,n)$ and the $\text{VIF}_M(k)$ for the concrete data of the afore mentioned example and for values of k varying between 0 and 0.3.
The $\text{VIF}_M(k)$ presents values lower than 1 for values of k higher than 0.09 while the values of VIF_{Sk} and $\text{VIF}_R(k,n)$ are always higher than 1. For values of k equal to 0.02 and 0.03 the conclusion will be very different. In these cases the VIF_{Sk} will be higher than 10, suggesting the possible existence of collinearity in the model, while the $\text{VIF}_M(k)$ will be lower than 10.
Furthermore, when k is equal to zero, the values of the three VIF expressions are equal. We can also appreciate that the VIF_{Sk} and the $\text{VIF}_R(k,n)$ are almost similar, since the difference between them is founded in the third decimal. However, the difference between the VIF_{Sk} and the $\text{VIF}_M(k)$ is larger, varying between 2 and 7.5.
As stated by O'Brien [36] *commonly a VIF of 10 or even one as low as 4 have been used as rules of thumb to indicate excessive or serious collinearity* [13, 25, 29, 31, 34, 35]. By considering these rules of thumb, the conclusion of this empirical application will be strongly different if we use the $\text{VIF}_R(k,n)$ instead of $\text{VIF}_M(k)$. For example, for k equal to 0.02, the $\text{VIF}_M(k)$ takes the value of 6.4199 and we could consider that the collinearity problem is solved. However, if we use the $\text{VIF}_R(k,n)$, the value of k that presents a VIF less than 10 will be 0.04. The same occurs if we consider that a VIF higher than 4 denotes the existence of collinearity problem. By using the $\text{VIF}_M(k)$, we obtain values of VIFs less than 4 from values of $k \geq 0.04$. However, when working with the proposed expression $\text{VIF}_R(k,n)$, we need to increase k until 0.14 to obtain a value of VIF less than 4.
On the other hand, we can appreciate that the $\text{VDF}_R(k)$ is almost zero from $k > 0.09$. Actually, we could state that the decrease of the variance in the ridge regression is explained by the $\text{VDF}_R(k)$ and its behavior. Thus, the $\text{VDF}_R(k)$ could explain why the variance of the parameter diminishes, even though the VIF (which measures the correlation between variables) is higher than 10 (see expression (35)). In our example,

we can see a strong correlation between the variables for $k = 0.2$ but the variance has diminished, due to the $\text{VDF}_R(k)$ which is equal to 0.000012.

8. Conclusions

We have shown that the expressions proposed in [29, 32] to calculate the Variance Inflation Factor (VIF) in ridge regression are not true VIFs. Marquardt [29] took the diagonal elements of the matrix $(\mathbf{X}'\mathbf{X} + k\mathbf{I})^{-1}\mathbf{X}'\mathbf{X}(\mathbf{X}'\mathbf{X} + k\mathbf{I})^{-1}$ as the variance inflation factor in the ridge regression ignoring the fact that this matrix is not in 'correlation form'. This uncorrected expression has been commonly repeated by several authors in the scientific existing literature leading to examples with values of VIF lower than 1 which evidently are not possible according to the definition of VIF. This paper has elucidated this issue.
We have presented the $\text{VIF}_R(k, n)$ as the VIF calculated from the matrix that generates the ridge estimators. We have proved that the limit of the $\text{VIF}_R(k, n)$, as n tends to infinity, coincides with the surrogate ridge regression VIF, VIF_{Sk}.
We have compared the $\text{VIF}_R(k, \infty)$ (or VIF_{Sk}) and the $\text{VIF}_R(k, n)$ with the $\text{VIF}_M(k)$. We have concluded that the VIF_{Sk} and the $\text{VIF}_R(k, n)$ are similar when $n > 30$ and that both increase when the correlation coefficient increases and the rest of variables keep constant. However, the $\text{VIF}_M(k)$ does not verify this logical behavior. On the contrary, it decreases when the correlation coefficient is close to one and the problems of collinearity are more serious (see Figure 5).
The contribution of this paper is illustrated with a numerical example previously applied [32, 33]. In this empirical application we show that the conclusions will vary greatly if we use the $\text{VIF}_R(k, n)$ instead of the $\text{VIF}_M(k)$. By following O'Brien [36] we have considered a value of VIF equal to 10 (or 4) as indicator of the existence of collinearity. Thus, the $\text{VIF}_M(k)$ will consider the problem is solved by providing values of VIF less than 10 (or 4) for values of k equal to 0.02 (or 0.04). However, if we used the proposed $\text{VIF}_R(k, n)$, these values of k will show that the collinearity problems still exist. It will be needed to increase the value of k until 0.04 (or 0.14) to obtain values of VIF less than 10 (or 4). Finally, taking into account that Mansfield and Billy [26] affirmed that *the existence of extreme pairwise correlations may be sufficient for detecting collinearity*, we consider that the $\text{VIF}_R(k, n)$ or the VIF_{Sk} should be applied to detect collinearity, instead of the $\text{VIF}_M(k)$.

Acknowledgements

We would like to thank referees for their detailed, encouraging and constructive reviews of our paper which clearly contributed to improving both the structure and the content of this manuscript.
This paper has been partially supported by the project 'Valoración de proyectos gubernamentales a largo plazo: obtención de la tasa social de descuento', reference: P09-SEJ-05404, Proyectos de Excelencia de la Junta de Andalucía and Fondos FEDER.

References

[1] J. E. Anderson, *Ridge regression of house value determinants*, J. Urban Econ. 9 (1981), pp. 286–297.
[2] B. Bruce and D. W. Hann, *Applications of ridge regression in forestry*, Forest Sci. 27 (1981), pp. 339–348.

[3] S. Das and S. Chatterjee, *Multicollinearity problem root cause diagnostic and way out*, Soc. Sci. Res. Network (2011), URL `http://ssrn.com/abstract=1830043`.
[4] J. Dias and J. Castro, *New multicollinearity indicators in linear regression models*, Int. Statist. Review 75 (2007), pp. 114–121.
[5] J. Dias and J. Castro, *The corrected vif*, J. Applied Statistics 38 (2011), pp. 1499–1507.
[6] L. Feng-Jeng, *Solving multicollinearity in the process of fitting regression model using the nested estimate procedure*, Qual. Quant. 42 (2008), pp. 417–426.
[7] J. Fox, *Applied regression analysis, linear models, and realted methods*, Thousand Oaks, CA: Sage, 1997.
[8] J. Fox and G. Monette, *Generalized collinearity diagnostics*, J. Amer. Statist. Assoc. 87 (1992), pp. 178– 183.
[9] C. B. García, J. García and J. Soto, *The raise method: An alternative procedure to estimate the parameters in presence of collinearity*, Qual. Quant. 45 (2010), pp. 403–423.
[10] R. Garg, *Ridge regression in the presence of multicollinearity*, Psychol. Rep. 54 (1984), pp. 559–566.
[11] W. H. Greene, *Econometric analysis*, 1993, 2nd ed.
[12] A. S. Hadi, *Ridge and Surrogate Ridge Regressions*, International Encyclopedia of Statistical Science, New York, NY: Springer, 2011.
[13] J. F. Hair, R. E. Anderson, R. L. Tatham and W. C. Black, *Multivariate data analysis*, 1995, 3rd ed.
[14] A. E. Hoerl, *Application of ridge analysis to regression problems*, Chemical Engineering Progress 58 (1962), pp. 215–223.
[15] A. E. Hoerl, *Ridge analysis*, Chemical Engineering Progress 60 (1964), pp. 67–77.
[16] A. E. Hoerl and R. W. Kennard, *Ridge regression: Applications to nonorthogonal problems*, Technometrics 12 (1970), pp. 69–82.
[17] A. E. Hoerl and R. W. Kennard, *Ridge regression: Biased estimation for nonorthogonal problems*, Technometrics 12 (1970), pp. 55–67.
[18] M. A. Ijomah and M. T. Nwakuya, *Ridge trace a boost to ridge regression estimate in the presence of multicollinearity*, Scientia Africana 10 (2011), pp. 42–47.
[19] M. Irfan, M. Javed and M. Ali Raza, *Comparison of shrinkage regression methods for remedy of multicollinearity problem*, Middle East J. Sci. Res. 14 (2013), pp. 570–579.
[20] N. Jamal and M. Q. Rind, *Ridge regression: A tool to forecast wheat area and production*, Pakistan J. Statist. Oper. Res. 3 (2007), pp. 125–134.
[21] D. R. Jensen and D. E. Ramirez, *Anomalies in the foundations of ridge regression*, International Statistical Review 76 (2008), pp. 89–105.
[22] D. R. Jensen and D. E. Ramirez, *Surrogate models in ill-conditioned systems*, Journal of Statistical Planning and Inference 140 (2010), pp. 2069–2077.
[23] D. R. Jensen and D. E. Ramirez, *Tracking MSE efficiencies in ridge regression*, Advances and Applications in Statistical Sciences 1 (2010), pp. 381–398.
[24] D. R. Jensen and D. E. Ramirez, *Revision: Variance inflation in regression*, Adv. Decis. Sci. 2013 (2013), pp. 1–15.
[25] P. A. Kennedy, *A guide to Econometrics*, MIT Press, 1992, 3rd ed.
[26] E. R. Mansfield and P. Billy, *Detecting multicollinearity*, Amer. Statist. 36 (1982), pp. 158–160.
[27] S. Mardikyan and E. Cetin, *Efficient choice of biasing constant for ridge regression*, Int. J. Contemp. Math. Sci. 3 (2008), pp. 527–536.
[28] C. Marinoiu, *A ridge regresion models fo the cracking proces*, Buletinul Universitati Petrol Gaze din Ploiesti. LXI (2009), pp. 65–70.
[29] D. W. Marquardt, *Generalized inverses, ridge regression, biased linear estimation and nonlinear estimation*, Technometrics 12 (1970), pp. 591–612.
[30] D. W. Marquardt and S. R. Snee, *Ridge regression in practice*, Amer. Statist. 29 (1975), pp. 3–20.
[31] R. L. Mason, R. F. Gunst and J. L. Hess, *Statistical design and analysis of experiments: Applications to engineering and sciences*, New York: Wiley, 1989.
[32] G. C. McDonald, *Tracing ridge regression coefficients*, Wiley Interdiscip. Rev.: Comput. Statist. 2 (2010), pp. 695–703.
[33] G. C. McDonald and R. C. Schwing, *Instabilities of regression estimates relating air pollution to mortality*, Technometrics 15 (1973), pp. 463–481.
[34] S. Menard, *Applied logistic regression analysis: Sage university Series on quantitative applications in the social sciences*, Thousand Oaks CA, Sage, 1995.
[35] J. Neter, W. Wasserman and K. M. H., *Applied linear regression models*, 1989, 3rd ed.
[36] R. M. O'Brien, *A caution regarding rules of thumb for variance inflation factors*, Qual. Quant. 41 (2007), pp. 673–690.
[37] B. Price, *Ridge regression: Application to nonexperimental data*, Psychol. Bull. 84 (1977), pp. 759–

766.
[38] H. Theil, *Principles of econometrics*, Wiley, New York, 1971.
[39] H. J. P. Timmermans, *Multiattribute shopping models and ridge regression analysis*, Environ. Plann. A 13 (1981), pp. 43–56.
[40] W. A. Williams, C. O. Qualset and S. Geng, *Ridge regession for extracting soybean yield factors*, Crop Sci. 19 (1979), pp. 869–873.
[41] J. Zhang and M. Ibrahim, *A simulation study on spss ridge regression and ordinary least squares regression procedures for multicollinearity data*, J. App. Stat. 32 (2005), pp. 571–588.

Tables

Table 1. Examples of VIF lower than one.

Authors	VIFs	k
Marquardt and Snee (1975), [30]	45.751, 0.826, 0.694, 0.510, 0.309	0, 0.1, 0.2, 0.4, 0.8
Price (1977), [37]	0.134	0.2
Williams et al. (1979), [40]	0.9	0.1
Timmermans (1981), [39]	0.736	0.5
Bruce and Hann(1981), [2]	VIF< 1	$k > 0.09$
Anderson (1981), [1]	0.75	1
Garg (1984), [10]	0.98, 0.94, 0.91, 0.88, 0.80	0.04, 0.05, 0.05, 0.07, 0.07
Jamal and Rind(2007), [20]	0.946	0.16
Mardikyan and Cetin (2008), [27]	0.993	0.021
Marinoiu (2009), [28]	0.656	0.124
McDonald (2010), [32]	0.823	0.1
Das and Chatterjee (2011), [3]	0.85	0.115
Ijomah and Nwakuya (2011), [18]	VIF< 1	$k > 0.055$
Irfan et al. (2013), [19]	0.994	0.02

Table 2. Explanation about the existence of a matrix $\mathbf{X}_k$ from which we have obtained $\mathbf{X}'\mathbf{X} + k\mathbf{I}$.

$\mathbf{X}$	$\rightarrow$	$\mathbf{X}'\mathbf{X}$	$\rightarrow$	$\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}$
$\mathbf{X}_k$	$\rightarrow$	$\mathbf{X}'\mathbf{X} + k\mathbf{I}$	$\rightarrow$	$\hat{\boldsymbol{\beta}}_R(k) = (\mathbf{X}'\mathbf{X} + k\mathbf{I})^{-1}\mathbf{X}'\mathbf{Y}$

Table 3. Summary of the behavior of the difference between the $\text{VIF}_R(k,n)$ and the VIF_{Sk}

	$\rho \in [-1; -0.9]$	$\rho \in [0.9; 1]$
$\frac{\partial \text{VIF}_R(k,n)}{\partial n}$	-	+
$\text{VIF}_R(k,n)$	Decreasing	Increasing
$\text{VIF}_R(k,n) - \text{VIF}_{Sk}$	> 0	< 0

Table 4. Values of VIF_{Sk}, $\text{VIF}(k, 60)$, $\text{VIF}_M(k)$ and $\text{VDF}_R(k)$ in the example proposed in [32].

k	VIF_{Sk}	$\text{VIF}(k, 60)$	$\text{VIF}_M(k)$	$\text{VIF}_{Sk} - \text{VIF}(k, 60)$	$\text{VIF}_{Sk} - \text{VIF}_M(k)$	$\text{VIF}(k, 60) - \text{VIF}_M(k)$	$\text{VDF}_R(k)$
0	31.5020	31.5020	31.5020	0.0000	0.0000	0.0000	1.000000
0.01	19.6763	19.6732	12.0838	0.0031	7.5925	7.5894	0.614130
0.02	14.4212	14.4167	6.4199	0.0045	8.0013	7.9968	0.445171
0.03	11.4514	11.4461	4.0253	0.0053	7.4261	7.4208	0.351512
0.04	9.5426	9.5369	2.7932	0.0058	6.7495	6.7437	0.292708
0.05	8.2127	8.2066	2.0763	0.0061	6.1363	6.1302	0.252816
0.06	7.2330	7.2266	1.6225	0.0064	5.6105	5.6041	0.224319
0.07	6.4814	6.4748	1.3168	0.0066	5.1646	5.1580	0.203166
0.08	5.8866	5.8799	1.1009	0.0067	4.7857	4.7790	0.187018
0.09	5.4043	5.3975	0.9426	0.0068	4.4617	4.4548	0.000017
0.1	5.0053	4.9984	0.8229	0.0069	4.1824	4.1754	0.000016
0.11	4.6698	4.6628	0.7301	0.0070	3.9397	3.9327	0.000016
0.12	4.3838	4.3767	0.6566	0.0071	3.7272	3.7201	0.000015
0.13	4.1371	4.1300	0.5973	0.0071	3.5399	3.5327	0.000014
0.14	3.9222	3.9150	0.5486	0.0072	3.3736	3.3664	0.000014
0.15	3.7333	3.7261	0.5082	0.0072	3.2252	3.2179	0.000014
0.16	3.5660	3.5587	0.4741	0.0073	3.0919	3.0846	0.000013
0.17	3.4167	3.4094	0.4450	0.0073	2.9717	2.9644	0.000013
0.18	3.2828	3.2755	0.4201	0.0073	2.8628	2.8554	0.000013
0.19	3.1620	3.1547	0.3984	0.0074	2.7636	2.7563	0.000013
0.2	3.0525	3.0451	0.3794	0.0074	2.6731	2.6657	0.000012
0.21	2.9527	2.9453	0.3627	0.0074	2.5900	2.5826	0.000012
0.22	2.8615	2.8541	0.3479	0.0074	2.5137	2.5062	0.000012
0.23	2.7778	2.7703	0.3346	0.0074	2.4432	2.4358	0.000012
0.24	2.7007	2.6932	0.3226	0.0075	2.3780	2.3706	0.000012
0.25	2.6294	2.6219	0.3118	0.0075	2.3176	2.3101	0.000012
0.26	2.5634	2.5559	0.3020	0.0075	2.2613	2.2539	0.000012
0.27	2.5020	2.4945	0.2931	0.0075	2.2089	2.2015	0.000012
0.28	2.4448	2.4373	0.2848	0.0075	2.1600	2.1525	0.000012
0.29	2.3915	2.3840	0.2773	0.0075	2.1142	2.1067	0.000012
0.3	2.3416	2.3340	0.2703	0.0075	2.0713	2.0638	0.000012

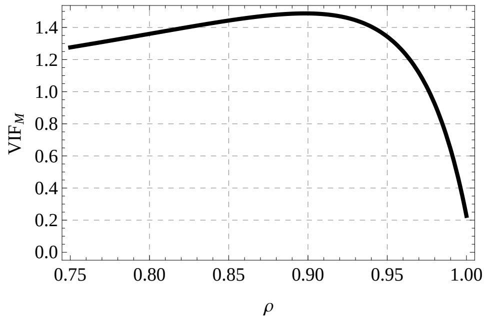
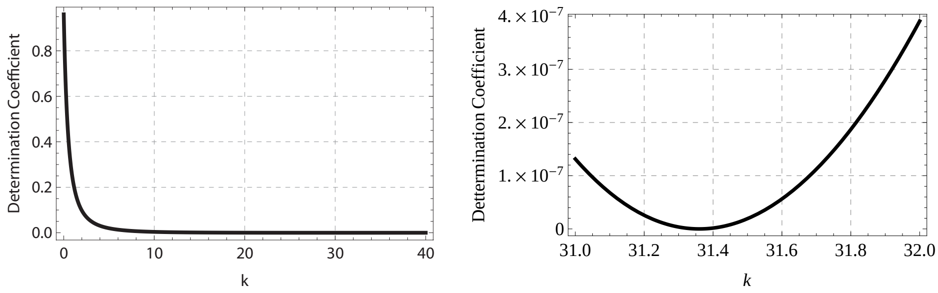


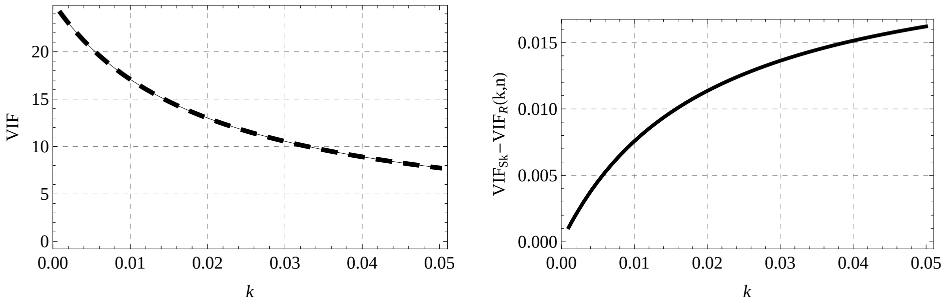
Figure 1. $\text{VIF}_M(k)$ for $k = 0.1$ and $0.75 \leq \rho \leq 1$.



(a) Evolution of determination coefficient (21) for $\rho = 0.98$, $n = 30$ and $0 \leq k \leq 40$.

(b) Detail when $31 \leq k \leq 32$ to observe the minimum value obtained when $k = 31.36$.

Figure 2. Representation of the determination coefficient presented in expression (21).



(a) Expression of the VIF_{Sk} (solid) and the $\mathrm{VIF}_R(k,n)$ (dashed).

(b) Representation of the difference between the VIF_{Sk} and the $\mathrm{VIF}_R(k,n)$.

Figure 3. VIF_{Sk} and $\mathrm{VIF}_R(k,n)$ for $n = 20$, $\rho = 0.98$ and $0 \leq k \leq 0.05$.

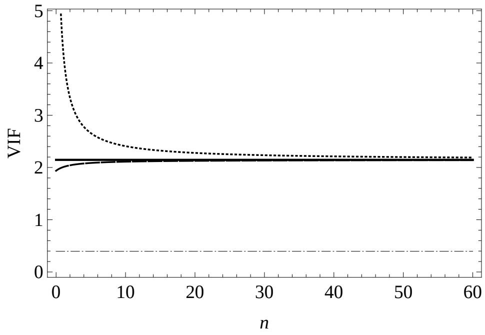


Figure 4. Comparison between the $\mathrm{VIF}_R(k,\infty)$ (solid), the $\mathrm{VIF}_R(k,n)$ for $\rho = 0.95$ (dashed), the $\mathrm{VIF}_R(k,n)$ for $\rho = -0.95$ (dotted) and the $\mathrm{VIF}_M(k)$ (dot-dashed), for $k = 0.3$.

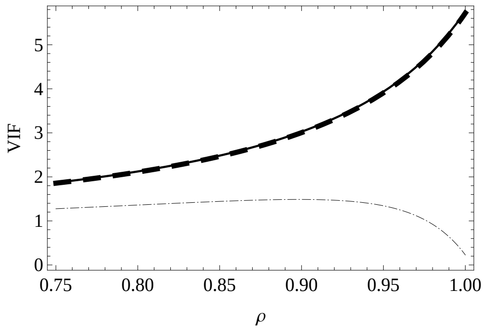


Figure 5. Comparison between the VIF_{Sk} (solid), the $\mathrm{VIF}_R(k,n)$ (dashed) and the $\mathrm{VIF}_M(k)$ (dot-dashed) for $0.75 \leq \rho \leq 1$, $n = 10$ and $k = 0.1$.