Journal of Modern Applied Statistical Methods

Volume 12 | Issue 2 Article 14

11-1-2013

How Good is Best? Multivariate Case of Ehrenberg-Weisberg Analysis of Residual Errors in Competing Regressions

Stan Lipovetsky
GfK Custom Research North America, Minneapolis, MN, stan.lipovetsky@gfk.com

Recommended Citation
Lipovetsky, Stan (2013) "How Good is Best? Multivariate Case of Ehrenberg-Weisberg Analysis of Residual Errors in Competing Regressions," *Journal of Modern Applied Statistical Methods*: Vol. 12 : Iss. 2 , Article 14.
DOI: 10.22237/jmasm/1383279180

Journal of Modern Applied Statistical Methods
November 2013, Vol. 12, No. 2, 242-255.

ISSN 1538 − 9472

How Good is Best? Multivariate Case of Ehrenberg-Weisberg Analysis of Residual Errors in Competing Regressions

Stan Lipovetsky
GfK Custom Research North America
Minneapolis, MN

A.S.C. Ehrenberg first noticed and S. Weisberg then formalized a property of pairwise regression to keep its quality almost at the same level of precision while the coefficients of the model could vary over a wide span of values. This paper generalizes the estimates of the percent change in the residual standard deviation to the case of competing multiple regressions. It shows that in contrast to the simple pairwise model, the coefficients of multiple regression can be changed over a wider range of the values including the opposite by signs coefficients. Consideration of these features facilitates better understanding the properties of regression and opens a possibility to modify the obtained regression coefficients into meaningful and interpretable values using additional criteria. Several competing modifications of the linear regression with interpretable coefficients are described and compared in the Ehrenberg-Weisberg approach.

Keywords: Pairwise and multiple regression, residual deviation change, Ehrenberg-Weisberg analysis

Introduction

In a fascinating work by A.S.C. Ehrenberg (1982) it was shown that the coefficients of pairwise regression can be varied over a wide span of values yet the modified model would still have a high quality of fit. Andrew Ehrenberg was a famous English statistician and marketing scientist recognized as the founder of probability models for consumer buying behavior (Ehrenberg, 1959, 1966, 1988; Fader and Hardie, 2009), and a prolific educator in statistics (for several examples, see Ehrenberg, 1981, 1983a,b). As Ehrenberg found, "The residuals from a least squares regression equation are hardly any smaller than those from many other possible lines" (1982, p. 364), and "markedly different equations give almost as good a fit as the least-squares regression equation itself" (1983a, p. 526). The

Dr. Lipovetsky is Senior Research Director at the Research Center for Excellence, GfK CRNA. Email him at: stan.lipovetsky@gfk.com.

technique considered by Ehrenberg was also described by S. Weisberg (1985, p. 68-70) in a convenient analytical form, thus, it will be called the Ehrenberg-Weisberg, or EW analysis.

EW analysis had been developed for pairwise models, but the current paper generalizes it to multiple regression where the results are even more interesting – in particular, it is even possible to change all the predictors' coefficients to the opposite sign, and still have almost the same precision of model fit. Such results show that the coefficients of linear regression can be adjusted by some additional criteria, where the coefficients become meaningful and the quality of the model stays high.

Regression modeling is widely used for statistical analysis and prediction in various problems of applied research. The main tool of regression modeling is the ordinary multiple linear least squares (OLS) regression which yields the best quality of data fit estimated by the minimum residual square error achieved by the aggregate of the predictors. However, OLS was not designed to obtain meaningful coefficients for individual predictors, and it is prone to multicollinearity effects which impact the coefficients' values and directions. Multicollinearity can make confidence intervals so wide that coefficients are incorrectly identified as insignificant, theoretically important variables receive negligible coefficients, or the coefficients have signs opposite to those of the corresponding pair correlations, so it is hardly possible to identify the individual predictors' importance in the regression (Grapentine, 1997; Mason and Perreault, 1991). Multicollinearity makes the covariance matrix of predictors close to singular, so its inversion yields inflated regression coefficients, pushing them to large values of both signs. It is difficult to use such an OLS solution for the analysis of key drivers, either by the coefficients or by the net effects (shares of the coefficient of multiple determination related to the predictors impact).

In the statistical literature and social sciences the effects of multicollinarity are explained by the so-called enhance, synergism, suppression, and masking effects among the predictors (Lipovetsky and Conklin, 2004). But such an explanation hardly helps to the interpretation and analysis of the regression results in applied research. For instance, in customer satisfaction studies in marketing research, the direction of the predictors' influence on the dependent variable is often known in advance. Suppose, the key drivers should all have a positive impact on overall satisfaction and it is evidenced by the pair correlations. But in OLS regression many coefficients turned out to be negative, so it is hardly possible to interpret the model and estimate the individual driver's importance. It is also difficult to use such a model for predicting a lift in the output because it is

not clear whether to increase or decrease a presumably useful variable if it has a negative sign in the model.

This article describes the features of EW analysis and its application to several modifications of multiple regression. One of those is the so-called Shapley value regression which is based on cooperative game theory used for finding the predictors' importance and then adjusting the regression coefficients via a nonlinear optimizing procedure. Another approach uses several modifications of the enhanced ridge regression technique to produce interpretable coefficients with a high overall quality of the model. A nonlinear parameterization of the coefficients of linear regression is also used in several forms to obtain sparse regression models with the features of interest. And finally, a model based on the elasticity criterion applied for building regression coefficients by data gradients is used for a comparison with OLS. In contrast to OLS, all the modified models are meaningful and easily interpretable, and have a quality of fit very close to the maximum defined by the OLS regression (for more detail on these models see (Lipovetsky and Conklin, 2001, 2010 a,b; Lipovetsky, 2009, 2010 a,b).

This paper is organized as follows: the next section describes the characteristics of EW for multiple and pairwise regressions, followed by a description of numerical simulations and a comparison of several modified regression solutions. A summary concludes the paper.

Ehrenberg-Weisberg Analysis

Consider briefly some relations from regression analysis needed for further development. For centered and normalized (by the standard deviations) dependent y_i and n design variables $x_{i1},...,x_{in}$ (i = 1,2,…,N – number of observations), a multiple linear regression model is:

$$y_i = b_1 x_{i1} + b_2 x_{i2} + ... + b_n x_{in} + e_i \tag{1}$$

where e_i denotes deviations from the model, and b are beta-coefficients of the standardized regression. In matrix form (1) can be represented as $y = Xb + e$, where y and e are the vectors of Nth order, and X is the matrix of N by n order. The least-squares objective is:

$$\begin{aligned} S^2 &= \| e \|^2 = \| y - Xb \|^2 = (y - Xb)'(y - Xb) \\ &= y'y - 2b'X'y + b'X'Xb = 1 - 2b'r + b'Rb, \end{aligned} \tag{2}$$

where for the standardized variables it is $y'y = 1$, the vector of pair correlations between y and each of n predictors x is $X'y = r$, and the matrix of pair correlations between the xs is $X'X = C$, and the prime denotes transposition. Minimizing by vector b yields the normal system of equations and its solution

$$Cb = r, \quad b = C^{-1}r, \tag{3}$$

where C^{-1} is inverted correlation matrix. Vector b (3) presents coefficients of the ordinary least squares, or OLS, regression. With OLS estimates b, the minimum residual sum of squares (2) and corresponding to it coefficient of multiple determination R^2 are defined as:

$$S^2 = 1 - b'r, \quad R^2 = 1 - S^2 = b'r = b'Cb \tag{4}$$

where r' is a transposed row-vector of correlations of x-s with y. The coefficient of multiple determination is always non-negative and less than one, its other properties are considered, for instance, in (Reisinger, 1997).

Next, describe EW, or Ehrenberg-Weisberg analysis deriving it from the beginning for the general case of multiple regression. Suppose each jth coefficient of regression b_j is changed with the term k_j, so the modified coefficients are

$$\tilde{b}_j = k_j b_j \tag{5}$$

or in the matrix form $\tilde{b} = diag(k)\,b$, where *diag(k)* is the diagonal matrix of terms k_j, and $\tilde{b}$ is the vector of modified coefficients of regression. With the new parameters $\tilde{b}$ the residual sum of squares (2) becomes:

$$\begin{aligned} \tilde{S}^2 &= \| y - X\tilde{b} \|^2 = (y - X\tilde{b})'(y - X\tilde{b}) \\ &= \left((y - Xb) - X(\tilde{b} - b)\right)' \left((y - Xb) - X(\tilde{b} - b)\right) \\ &= (y - Xb)'(y - Xb) - 2(\tilde{b} - b)'X'(y - Xb) + (\tilde{b} - b)'X'X(\tilde{b} - b) \end{aligned} \tag{6}$$

Taking (3) into account, the middle item in (6) equals zero because $X'(y - Xb) = r - Cb = 0$, so (6) can be reduced to:

$$\tilde{S}^2 = S^2 + (\tilde{b} - b)'C(\tilde{b} - b) = S^2 + b'\left(diag(k) - I\right)C\left(diag(k) - I\right)b \qquad (7)$$

In a simple case when all coefficients are changed by the same term $k_j = k$, taking (4) into account, the expression (7) can be reduced to:

$$\tilde{S}^2 = S^2 + (1-k)^2 b'Cb = S^2 + (1-k)^2 R^2 \qquad (8)$$

Dividing (8) by S^2 and using (4) yields the relation:

$$\frac{\tilde{S}^2}{S^2} = 1 + (1-k)^2 \frac{R^2}{1-R^2} \qquad (9)$$

For the simple pairwise regression by only one predictor ($n = 1$ in (1)) this formula coincides with the one obtained by Ehrenberg and Weisberg up to the change of the multiple correlation R to the pair correlation, $R^2 = r^2$. Taking the square root of (9) produces a quotient of the standard errors of OLS to the modified model expressed as:

$$\left(\frac{\tilde{S}^2}{S^2}\right)^{1/2} = \left(1 + (1-k)^2 \frac{R^2}{1-R^2}\right)^{1/2} \qquad (10)$$

It is the formula given in Weisberg (1985, p. 69) for the pairwise model with $R^2 = r^2$. Because the OLS solution has minimum standard error, (10) can be represented as

$$\left(1 + (1-k)^2 \frac{R^2}{1-R^2}\right)^{1/2} = 1 + d \qquad (11)$$

where $d > 0$ denotes the relative difference of the modified model's and OLS standard errors.

If d is assumed to be at a desirable level, for example, 5% or 10% , then it is possible find the range of k values for which the regression coefficient can be changed but the standard error will be kept within a d% increase from the minimum OLS standard error value:

$$\left(\tilde{S}^2\right)^{1/2} / \left(S^2\right)^{1/2} \le 1+d \tag{12}$$

For this aim, the inequality for k is solved as:

$$\left(1+(1-k)^2 \frac{R^2}{1-R^2}\right)^{1/2} \le 1+d \tag{13}$$

and the solution is:

$$1-\left((2d+d^2)\frac{1-R^2}{R^2}\right)^{1/2} \le k \le 1+\left((2d+d^2)\frac{1-R^2}{R^2}\right)^{1/2} \tag{14}$$

So for regression coefficient change with the term k (5) in the range (14) the inequality (12) is satisfied. With R^2 close to 1 the range (14) is narrow, but with small R^2 the modified coefficient of regression (5) can vary in the wide span without changing much of the residual error. For example, if d = 5%, the span (14) is given by the inequalities:

$$1-0.32\left(\frac{1-R^2}{R^2}\right)^{1/2} \le k \le 1+0.32\left(\frac{1-R^2}{R^2}\right)^{1/2} \tag{15}$$

or the span for the regression coefficient keeping the residual error in the limit of d = 10% is:

$$1-0.46\left(\frac{1-R^2}{R^2}\right)^{1/2} \le k \le 1+0.46\left(\frac{1-R^2}{R^2}\right)^{1/2} \tag{16}$$

It could seem that for small R^2 (for instance if |R|<0.3 in (15), or |R|<.4 in (16)) k can even be negative, so the regression changes its direction. However, for the pairwise regression it is not so, and it is not so for the multiple regression if all parameters of change are constant, $k_j = k$. Indeed, using the coefficients of multiple determination of OLS (4) and of the modified regression $\tilde{R}^2 = 1-\tilde{S}^2$, the equality (8) is represented as:

$$1-\tilde{R}^2 = 1-R^2+(1-k)^2R^2 \tag{17}$$

which reduces to:

$$\tilde{R}^2 = k(2-k)R^2 \tag{18}$$

To keep $\tilde{R}^2 \geq 0$, the values of *k* should belong to the range $0 \leq k \leq 2$. Thus, *k* in (14)-(16) cannot become negative for the pairwise regression, and the same holds for multiple regression in the case where all *k* are equal. The sufficient condition to have $0 \leq k \leq 2$, so $\tilde{R}^2 \geq 0$, is to keep the square roots in (14) less than one:

$$\left((2d+d^2)\frac{1-R^2}{R^2}\right)^{1/2} \leq 1 \tag{19}$$

which can be represented more concisely as follows:

$$(1+d)^2(1-R^2) \leq 1 \tag{20}$$

Thus, for a given value of R^2 the percent *d* satisfying the condition (20) which guarantees the modified $\tilde{R}^2$ to be positive should be chosen.

Continuing with EW description for multiple regression in a general case where different parameters of change are assigned to each coefficient, similar to the transformation of (8) to (9), the general expression (7) can be presented in explicit form as:

$$\frac{\tilde{S}^2}{S^2} = 1+\frac{1}{1-R^2}\left(\sum_{j=1}^{n}(1-k_j)^2b_j^2 + 2\sum_{j>q}^{n}(1-k_j)(1-k_q)b_jb_qr_{jq}\right) \tag{21}$$

where r_{jq} are the pair correlations between x_j and x_q. The terms with $1 - k_j$ in (21) modify the inputs from b_j^2 (the so-called pure net-effects of each predictor) and from $b_jb_qr_{jq}$ (the so-called mixed net-effects of the predictors) into the coefficient of multiple determination. If only one coefficient of regression is changed, say, $k_j \neq 1$, and all the others are kept intact ($k = 1$) then the ratio (21) reduces to:

$$\frac{\tilde{S}^2}{S^2} = 1 + (1-k_j)^2 \frac{b_j^2}{1-R^2} \tag{22}$$

From (22) with the net effect of the *j*th predictor in the numerator, it is easy to derive the relations (10)-(14) for considering a model with only one modified coefficient. But a general case of different changes for all the coefficients (21) can be studied in numerical simulations.

Numerical Simulation and Examples

Consider the case of two predictors, $n = 2$, trying several values of pairwise correlations r_{y1} and r_{y2} of y with x_1 and x_2, and the r_{12} correlation between two predictors taken within the allowed range of the values:

$$r_{y1}r_{y2} - \sqrt{(1-r_{y1}^2)(1-r_{y2}^2)} \;\le\; r_{12} \;\le\; r_{y1}r_{y2} + \sqrt{(1-r_{y1}^2)(1-r_{y2}^2)} \tag{23}$$

Table 1. Numerical simulation with various *k*: pair correlations, OLS regressions, *k*-terms, modified regressions, and residual STD change.

Pair correlations			OLS regression			Terms of change		Modified regression			STD change
r_{y1}	r_{y2}	r_{12}	b_1	b_2	R^2	k_1	k_2	$\tilde{b}_1$	$\tilde{b}_2$	$\tilde{R}^2$	d
-0.75	0.75	-0.900	-0.395	0.395	0.592	-0.1	2.0	0.039	0.789	0.556	0.043
-0.75	0.75	-0.900	-0.395	0.395	0.592	0.1	2.0	-0.039	0.789	0.562	0.036
-0.75	0.75	-0.900	-0.395	0.395	0.592	0.5	2.0	-0.197	0.789	0.538	0.065
-0.75	0.75	-0.813	-0.414	0.414	0.621	2.0	-0.1	-0.828	-0.041	0.548	0.091
-0.75	0.75	-0.813	-0.414	0.414	0.621	2.0	0.1	-0.828	0.041	0.561	0.076
-0.75	0.75	-0.813	-0.414	0.414	0.621	2.0	0.5	-0.828	0.207	0.546	0.094
-0.50	0.50	-0.150	-0.435	0.435	0.435	0.5	0.5	-0.217	0.217	0.326	0.092
0.10	0.50	0.739	-0.595	0.940	0.410	0.5	0.5	-0.297	0.470	0.308	0.084
0.50	0.50	0.100	0.455	0.455	0.455	0.5	0.5	0.227	0.227	0.341	0.099
0.50	0.75	0.490	0.175	0.664	0.586	2.0	0.5	0.349	0.332	0.502	0.097
0.50	0.75	0.604	0.074	0.705	0.566	5.0	0.5	0.369	0.353	0.480	0.094
0.50	0.75	0.719	-0.081	0.808	0.566	-2.0	0.5	0.161	0.404	0.484	0.090
0.75	0.75	0.825	0.411	0.411	0.616	2.0	-0.1	0.822	-0.041	0.550	0.083
0.75	0.75	0.825	0.411	0.411	0.616	2.0	0.1	0.822	0.041	0.562	0.069
0.75	0.75	0.825	0.411	0.411	0.616	2.0	0.5	0.822	0.205	0.545	0.090

Table 1 presents the sets of these three correlations in the first three columns, and in the next three columns there are OLS beta-coefficients of regression (1) and the coefficient of multiple determination R^2 (4). The terms k_j for the modified coefficients are given in the two middle columns of Table 1, then there are the modified coefficients themselves (5), and the corresponding modified coefficient $\tilde{R}^2 = 1 - \tilde{S}^2$ of multiple determination. The last column of Table 1 presents the relative change of the residual standard deviation (STD), which can be expressed via (12) and (21) as follows:

$$d = \sqrt{1 + \frac{(1-k_1)^2 b_1^2 + (1-k_2)^2 b_2^2 + 2(1-k_1)(1-k_2) b_1 b_2 r_{12}}{1-R^2}} - 1 \qquad (24)$$

As shown in Table 1, the change of coefficients can be very noticeable but the change in STD is below 10% of the precision in all fifteen examples given in rows.

Table 2. Numerical simulation with negative *k*: pair correlations, OLS regressions, *k*-terms, modified regressions, and residual STD change.

Pair correlations			OLS regression			Terms of change		Modified regression			STD change
r_{y1}	r_{y2}	r_{12}	b_1	b_2	R^2	$k_1<0$	$k_2<0$	$\tilde{b}_1$	$\tilde{b}_2$	$\tilde{R}^2$	d
-0.50	-0.25	0.628	-0.566	0.106	0.257	-0.1	-2.0	0.057	-0.212	0.016	0.151
-0.50	-0.25	0.628	-0.566	0.106	0.257	-0.1	-2.0	0.028	-0.212	0.039	0.137
-0.50	-0.25	0.628	-0.566	0.106	0.257	-0.1	-1.0	0.028	-0.106	0.016	0.150
-0.50	-0.25	0.796	-0.821	0.403	0.310	-0.1	-1.0	0.082	-0.403	0.003	0.202
-0.50	-0.25	0.796	-0.821	0.403	0.310	-0.1	-1.0	0.041	-0.403	0.023	0.190
-0.50	-0.25	0.796	-0.821	0.403	0.310	-0.1	-0.5	0.041	-0.202	0.031	0.185
-0.25	-0.10	0.603	-0.298	0.080	0.067	-0.1	-2.0	0.015	-0.160	0.002	0.034
-0.25	-0.10	0.603	-0.298	0.080	0.067	-0.1	-1.0	0.015	-0.080	0.003	0.033
-0.25	-0.10	0.796	-0.465	0.270	0.089	-0.1	-0.5	0.023	-0.135	0.002	0.047
0.50	0.75	0.719	-0.081	0.808	0.566	-5.0	-0.1	0.404	-0.040	0.202	0.356
0.50	0.75	0.719	-0.081	0.808	0.566	-2.0	-0.1	0.161	-0.081	0.026	0.497
0.50	0.75	0.719	-0.081	0.808	0.566	-2.0	-0.1	0.161	-0.040	0.082	0.453
0.50	0.75	0.833	-0.409	1.091	0.614	-2.0	-0.1	0.817	-0.055	0.139	0.493
0.50	0.75	0.833	-0.409	1.091	0.614	-1.0	-0.1	0.409	-0.109	0.140	0.491
0.50	0.75	0.833	-0.409	1.091	0.614	-1.0	-0.1	0.409	-0.055	0.194	0.444

Table 2 is organized as Table 1 but it contains both the *k*-terms of negative sign, $k_1 < 0$ and $k_2 < 0$, so the direction of *y*'s connections with the predictors is flipped. Table 2 shows that although the direction of the model coefficients can be changed, the quality of such models is not high, and the precision of STD change could be low too. As can be expected, a model could receive the opposite signs of the coefficients and keep about the same quality of fit mostly in the cases of weak statistical relationships similar to those considered in (Langford et al., 2001).

As it was discussed in the introduction, because of the effects of multicollinearity the coefficients of regression can be found in a wide range of the values of both signs. It can be shown in a simple example of the model with two predictors where the beta-coefficients of regression are defined as follows:

$$b_1 = \frac{r_{y1} - r_{y2}r_{12}}{1 - r_{12}^2}, \quad b_2 = \frac{r_{y2} - r_{y1}r_{12}}{1 - r_{12}^2} \tag{25}$$

Suppose all correlations are positive, and x_1 is strongly correlated with x_2, so r_{12} is close to 1. Then the numerators in the coefficients (25) become close to $r_{y1} - r_{y2}$ and $r_{y1} - r_{y2}$, respectively, so of opposite signs. At the same time the denominator $1 - r_{12}^2$ is close to zero, so b_1 and b_2 become big by the absolute value and of opposite signs. It is effect of inflation under multicollinearity, and changing directions of the connection from positive pairwise to opposite by sign in multiple regression. Using various methods of regularization, mentioned in the introduction, meaningful regression coefficients can be obtained. And the EW relative change of the residual standard deviations can be used for comparison of the several competing regression models and checking how far are the residual errors from their OLS minimum value.

Consider a numerical example where several regressions were tried by the data on various cars' characteristics given in (Chambers and Hastie, 1992; and also available in *S-PLUS'2000*, 1999, as "car.all" data). The data describes dimensions and mechanical specifications supplied by the manufacturers and measured by Consumer Reports. The variables are: *y* – Price of a car, US$K; x_1 – Weight, pounds; x_2 – Length overall, inches; x_3 – Wheel base length, inches; x_4 – Width, inches; x_5 – Front Leg Room maximum, inches; x_6 – Front Shoulder room, inches; x_7 – Turning circle radius, feet; x_8 – Displacement of the engine, cubic inches; x_9 – HP, the net horsepower; x_{10} – Tank fuel refill capacity, gallons. The cars' price is estimated in the regression model by the dimensions and specifications variables.

Table 3. Cars example: correlations and several regressions.

Name	variable	r_{yx}	OLS	SV	Grad	RR	RE2	RE3	Exp	Multin
Weight	x_1	0.653	0.278	0.129	0.101	0.053	0.105	0.116	0.088	0.000
Length	x_2	0.533	0.225	0.072	0.083	0.039	0.056	0.062	0.066	0.099
Wheel .base	x_3	0.496	-0.085	0.043	0.077	0.032	0.034	0.038	0.000	0.000
Width	x_4	0.478	-0.144	0.047	0.074	0.029	0.024	0.026	0.000	0.000
Frt.Leg .Room	x_5	0.567	0.245	0.140	0.088	0.063	0.129	0.143	0.258	0.248
Frt.Shld	x_6	0.371	-0.060	0.012	0.057	0.017	0.006	0.007	0.000	0.000
Turning	x_7	0.378	-0.199	0.022	0.059	0.017	0.003	0.003	0.000	0.000
Disp.	x_8	0.642	0.101	0.110	0.100	0.053	0.097	0.107	0.000	0.000
HP	x_9	0.783	0.409	0.191	0.121	0.082	0.293	0.323	0.512	0.528
Tank	x_{10}	0.657	0.160	0.114	0.102	0.056	0.116	0.128	0.085	0.125
	R^2		0.722	0.596	0.503	0.409	0.637	0.645	0.695	0.694
	d			0.205	0.337	0.458	0.143	0.130	0.047	0.049

Table 3 in the first and second numerical columns presents the pair correlations r_{yx} of y with x, and the OLS beta-coefficients (1). All correlations are positive, but four of the ten variables have negative coefficients in the multiple OLS regression, although it has a good coefficient of multiple determination $R^2 = 0.722$. The next seven columns in Table 3 present several modified solutions referred to in the introduction: *SV* – Shapley value model, *Grad* – the model constructed by the data gradients; *RR* – the regular ridge regression, *RE2* and *RE3* – two kinds of the ridge enhanced models, *Exp* and *Multin* – the model with exponential and multinomial-logit parameterization of the coefficients of multiple linear regression. Below each model, its coefficient of multiple determination is shown, together with the EW relative change characteristic of the residual standard deviation *d*.

All the modified models have non-negative coefficients of regression, and their coefficients R^2 are slightly less than the maximum R^2 of OLS. But the more sensitive characteristic of *d* indicates rather clearly that *RR* and *Grad* models are fair, the *SV* and both *RE* models are good, and the *Exp* and *Mult* models give the best variants with less than 5% of the difference in standard deviations. As had been shown with more detail in (Lipovetsky, 2009, 2010a,b), the enhanced and adjusted ridge models systematically outperform regular ridge regression, and

special parameterization techniques produce nonnegative coefficients with a clear, sparse structure in the two last approaches. As an additional useful feature of the *Mult* model, the total of the beta-coefficients equals exactly one, so the coefficients equivalent to the shares of the predictors' impact on the dependent variable. However, if it is desirable to keep and compare all the variables in the model then the *SV* and *ridge* regressions should be used, and the *Grad* model is preferable for an express analysis when no special software is available.

Summary

A modified least squares regression can have better interpretable coefficients and practically the same quality of fit, which can be estimated by the characteristic of the relative change in the residual standard deviation. This paper develops the Ehrenberg-Weisberg estimation of the characteristic of relative change in the residual standard deviation for pair regression to the general case of multiple regression. It shows that the coefficients of ordinary least-squares can be changed over a wide range of values, including the opposite sign, and the quality of fit can still be at an acceptable level. This estimation is applied for a comparison of several regressions with the ordinary least squares model, to identify the modified regressions with interpretable coefficients and good quality of fit. The obtained results help provide a better understanding of the properties of multiple regression, and are useful for theoretical consideration and practical applications of regression modeling and analysis.

References

Chambers, J.M., and Hastie, T.J. (1992). *Statistical Models in S*. Pacific Grove, CA: Wadsworth and Brooks.

Ehrenberg, A.S.C. (1959). The Pattern of Consumer Purchases. *Applied Statistics*, *8*, 26–41.

Ehrenberg, A.S.C. (1966). Laws in Marketing: A Tail-Piece. *Journal of the Royal Statistical Society, Series C, 15*, 257-267.

Ehrenberg, A.S.C. (1981). The Problem of Numeracy. *The American Statistician, 35*, 67-71.

Ehrenberg, A.S.C. (1982). How Good is Best? *Journal of the Royal Statistical Society, Series A, 145*, 364-366.

Ehrenberg, A.S.C. (1983a). Lawlike relationships. In: *Encyclopedia of Statistical Sciences* (N. L. Johnson & S. Kotz, Eds.), *4*, 523-528, NY: Wiley.

Ehrenberg, A.S.C. (1983b). Deriving the Least Squares Regression Equation. *The American Statistician*, *37*, 232.

Ehrenberg, A. S. C. (1988). *Repeat-Buying*. 2nd ed. London: Griffin.

Fader, P.S., and Hardie, B.G.S. (2009). Probability Models for Customer-Base Analysis. *J. of Interactive Marketing, 23*, 61-69.

Grapentine, T. (1997). Managing Multicollinearity. *Marketing Research*, *9,* 11-21.

Langford, E., Schwertman, N., and Owens, M. (2001). Is the Property of Being Positively Correlated Transitive?. *The American Statistician, 55,* 322-325.

Lipovetsky, S. (2009). Linear Regression with Special Coefficient Features Attained via Parameterization in Exponential, Logistic, and Multinomial-Logit Forms. *Mathematical and Computer Modelling*, *49*, 1427-1435.

Lipovetsky, S. (2010a). Enhanced Ridge Regressions. *Mathematical and Computer Modelling, 51*, 338-348.

Lipovetsky, S. (2010b). Meaningful Regression Coefficients Built by Data Gradients. *Advances in Adaptive Data Analysis*, *2*, 451–462.

Lipovetsky, S., and Conklin, M. (2001). Analysis of Regression in Game Theory Approach. *Applied Stochastic Models in Business and Industry*, *17,* 319-330.

Lipovetsky, S., and Conklin, M. (2004). Enhance-Synergism and Suppression Effects in Multiple Regression. *International Journal of Mathematical Education in Science and Technology, 35*, 391-402.

Lipovetsky, S., and Conklin, M. (2010a). Reply to the paper 'Do not adjust coefficients in Shapley value regression'. *Applied Stochastic Models in Business and Industry*, *26*, 203-204.

Lipovetsky, S., and Conklin, M. (2010b). Meaningful Regression Analysis in Adjusted Coefficients Shapley Value Model. *Model Assisted Statistics and Applications, 5*, 251-264.

Mason, C.H., and Perreault, W.D. (1991). Collinearity, power, and interpretation of multiple regression analysis. *Journal of Marketing Research*, *28,* 268-280.

Reisinger, H. (1997) The impact of research designs on R^2 in linear regression models: an exploratory meta-analysis. *Journal of Empirical Generalisations in Marketing Science, 2*, 1-12.

S-PLUS'2000 (1999). Seattle, WA: MathSoft.
Weisberg, S. (1985). *Applied Linear Regression*, 2nd ed. New York: Wiley.