2017

Prediction of Percent Change in Linear Regression by Correlated Variables

Stan Lipovetsky

GfK North America, Minneapolis, MN, stan.lipovetsky@gfk.com

Prediction of Percent Change in Linear Regression by Correlated Variables

Stan Lipovetsky
GfK North America
Minneapolis, MN

Multiple linear regression can be applied for predicting an individual value of dependent variable *y* by the given values of independent variables *x*. But it is not immediately clear how to estimate percent change in *y* due to changes in predictors, especially when those are correlated. This work considers several approaches to this problem, including its formulation via predictors adjusted by their correlation structure. Ordinary least squares regression is used, together with Shapley value regression and another model based on solving some system of differential equations. Numerical estimations performed for a real marketing research data demonstrate meaningful results. The considered techniques can be very useful in practical estimations of the percent change of dependent variable by the change in predictors.

Keywords: Multiple regression, percent change in outcome, predictors percent change, correlated structure, OLS, Shapley value regression, differential equations

Introduction

Ordinary least squares (OLS) multiple linear regression is one of the main tools of statistical modeling widely used for estimation the dependent variable (DV) value by the values of independent variables (IVs), or predictors. In applied studies, often the need is to estimate not a DV value itself but a percent change in the outcome due to percent changes in IVs. Several problems should be solved in such estimations, beginning from finding a good criterion for evaluation of the percent change in the outcome. The mean values of DV and IVs satisfy a linear regression model with intercept, and this relation is employed for measuring percentage change in the output due to changes in the input variables.

Also, the predictors are only called independent in contrast to the dependent variable, but IVs in a sampled data are always correlated and not statistically

Dr. Stan Lipovetsky is the Senior Research Director in the Marketing Sciences department. Email him at: stan.lipovetsky@gfk.com.

independent. So the correlation structure should be accounted in finding the IVs' values for adequate prediction by regression. And there is a problem of multicollinearity among IVs which produces regressions with inflated values of coefficients, yields their signs opposite to the signs of the pair correlations, makes theoretically important variables to get small coefficients, causes a reduction in statistical power, and leads to wider confidence intervals for the coefficients so they could be incorrectly identified as being insignificant (Grapentine, 1997; Mason & Perreault, 1991). For instance, in marketing research, it is often known in advance that the influence of each IV on the DV of customer satisfaction should be positive, and it is supported by the pair correlations. But in the OLS regression many coefficients can occur to be negative, so for predicting a change in the output should we increase or decrease a presumably beneficial variable which, however, has a negative coefficient in the model?

To overcome deficiencies of multicollinearity and produce meaningful regression parameters, various modifications of OLS have been developed. Among those are: ridge regressions (Hoerl & Kennard, 1970; Lipovetsky, 2010), Shapley value regression (SVR) based on cooperative game theory used for finding predictors' importance and adjusting the regression coefficients (Shapley, 1953; Roth, 1988; Lipovetsky & Conklin, 2001), nonlinear parameterization of linear regression coefficients by multinomial shares, using elasticity criterion for building regression coefficients by data gradients, Gibson-Johnson and Johnson indices of predictor importance, and other techniques (for more detail and references within: Gibson, 1962; R. Johnson, 1966; J. Johnson, 2000; Lipovetsky, 2013; Lipovetsky & Conklin, 2014). Those techniques produce very similar models so we can employ one of them, SVR, as the model with interpretable coefficients.

Several approaches to the problem of estimating a percent change in the DV due to the percent changes in the IVs are considered here. OLS and SVR are used as the models for predicting change in the outcome for a given set of the predictors' values. These values are adjusted due to the structure of correlations among the predictors. For this aim, the mutual regressions of each predictor by all the other predictors are used. In another approach, a system of linear differential equations is considered as well.

Change Estimation by Multiple Regression

Consider several main relations of the OLS regression that will be needed further. A multiple linear regression can be presented as a model:

$$y_i = a_0 + a_1 x_{i1} + \ldots + a_n x_{in} + \varepsilon_i \equiv \hat{y}_i + \varepsilon_i \tag{1}$$

where x_{ij} and y_i are i^{th} observations ($i = 1, \ldots, N$ – number of observations) by each jth IV ($j = 1, \ldots, n$ – number of variables) and by the DV, a_j are the coefficients of regression, a_0 is its intercept, $\hat{y}_i$ denotes theoretical linear aggregate of the predictors, and ε_i are the deviations from the theoretical relationship. Least squares (LS) objective for finding coefficients consists in minimization of the total of squared deviations:

$$S^2 = \sum_{i=1}^{N} \varepsilon_i^2 = \sum_{i=1}^{N} \left(y_i - \hat{y}_i \right)^2 = \sum_{i=1}^{N} \left(y_i - a_0 - a_1 x_{i1} - \ldots - a_n x_{in} \right)^2 \tag{2}$$

Derivatives $\partial S^2 / \partial a_j = 0$ yield the normal system of equations for finding parameters of the model. Such a derivative by a_0 equaled zero produces the expression:

$$\overline{y} = a_0 + a_1 \overline{x}_1 + \ldots + a_n \overline{x}_n \tag{3}$$

where a bar above variables denotes their mean values. This relation is used to calculate the intercept when other coefficients of regression are found in the solution of the normal system which can be expressed in the matrix form as follows:

$$\mathbf{a} = \mathbf{C}_{xx}^{-1} \mathbf{c}_{xy} \tag{4}$$

where $\mathbf{a}$ denotes the vector of coefficients $a_1, \ldots, a_n$ of regression, $\mathbf{C}_{xx}$ and $\mathbf{C}_{xx}^{-1}$ are the n-by-n covariance matrix between xs and its inverse matrix, respectively, and $\mathbf{c}_{xy}$ is the nth order vector of covariance between xs and y. The results (3)-(4) present the OLS regression parameters.

Suppose the mean level of each driver $\overline{x}_j$ in the model can be increased by different portions f_j ($100 f_j$ percent), so the absolute change (denoted by d) and new values (denoted by prime) can be written as follows:

$$d\overline{x}_j = \overline{x}_j f_j, \quad \overline{x}'_j = \overline{x}_j + d\overline{x}_j = \overline{x}_j \left(1 + f_j \right) \tag{5}$$

Multiplying the parameters (4) by the vector of new values $\overline{\mathbf{x}}'$ and adding the intercept, we find a predicted DV value by the regression model. Subtracting (3) from it yields the similar relation for the changes in mean levels:

$$dy = a_1 d\overline{x}_1 + \ldots + a_n d\overline{x}_n \tag{6}$$

Dividing both sides of (6) by the mean level of y, and dividing each change in x_j by the corresponding mean value, we transform (6) to the following relation

$$\frac{d\overline{y}}{\overline{y}} = \left(a_1 \frac{\overline{x}_1}{\overline{y}} \right)\frac{d\overline{x}_1}{\overline{x}_1} + \ldots + \left(a_n \frac{\overline{x}_n}{\overline{y}} \right)\frac{d\overline{x}_n}{\overline{x}_n} = \left(a_1 \frac{\overline{x}_1}{\overline{y}} \right)f_1 + \ldots + \left(a_n \frac{\overline{x}_n}{\overline{y}} \right)f_n \tag{7}$$

which presents the relative (%) change in the DV outcome via % changes f_j in the IVs (5) and regression coefficients modified to the form $a_j \overline{x}_j / \overline{y}$.

Instead of OLS regressions (4) it is possible to use Shapley value regression (Lipovetsky & Conklin, 2001). Coefficients of SVR have a more adequate meaning (similar to elasticity but in absolute changes) as a change in the output due to the unit change in each predictor holding other predictors constant. The modified coefficients in (7) also have a clear interpretation if considered via the SVR parameters.

Adjusting Independent Variables by Correlation Structure

Suppose the percentage changes f_j for xs are given so they can be used in (5) for finding the new values $\overline{x}'_j$. However, these values for prediction should be adjusted due to the correlations among the predictors given in the matrix $\mathbf{C}_{xx}$. A convenient way to such an adjustment is as follows: The diagonal elements of inverted correlation matrix $\mathbf{C}_{xx}^{-1}$ used in (4) (called variance inflation factors, VIF) equal the reciprocal values of the residual sums of squares in the regressions of each variable x_j by the rest of IVs,

$$\text{VIF}_j = \left(\mathbf{C}_{xx}^{-1} \right)_{jj} = 1 / \left(1 - R_{j.}^2 \right) \tag{8}$$

where $R_{j.}^2$ are the coefficients of multiple determination in the models of each x_j by the other xs. These models in the so-called Yule's notations are:

$$x_j = a_{j0} + a_{j1.}x_1 + a_{j2.}x_2 + \ldots + a_{j,j-1.}x_{j-1} + a_{j,j+1.}x_j + \ldots + a_{jn.}x_n \qquad (9)$$

where $a_{jk.}$ denotes a parameter of jth regression by k^{th} variable among all $n-1$ other xs (Kendall & Stuart, 1973; Lipovetsky & Conklin, 2004). The non-diagonal elements in any j^{th} row of $\mathbf{C}_{xx}^{-1}$ divided by the diagonal element in the same row and taken with opposite signs coincide with coefficients of regression x_j by all other xs (9) that can be presented as the following matrix $\mathbf{A}$:

$$\mathbf{A} \equiv \text{diag}\left(\mathbf{C}_{xx}^{-1}\right)^{-1}\mathbf{C}_{xx}^{-1} = \begin{pmatrix} 1 & -a_{12.} & -a_{13.} & \ldots & -a_{1n.} \\ -a_{21.} & 1 & -a_{23.} & \ldots & -a_{2n.} \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ -a_{n1.} & -a_{n2.} & -a_{n3.} & \ldots & 1 \end{pmatrix} \qquad (10)$$

The product $\mathbf{A}\bar{\mathbf{x}}$ of the matrix (10) and the vector of mean values $\bar{\mathbf{x}}$ of all predictors coincides with the vector of intercepts $\mathbf{a}_0$ for the mutual regressions (9):

$$a_{j0} = \bar{x}_j - \left(a_{j1.}\bar{x}_1 + a_{j2.}\bar{x}_2 + \ldots + a_{j,j-1.}\bar{x}_{j-1} + a_{j,j+1.}\bar{x}_j + \ldots + a_{jn.}\bar{x}_n\right) \qquad (11)$$

Multiplying (10) by the vector of the new values $\bar{\mathbf{x}}'$ (with elements $\bar{x}'_j$ (5)) and subtracting the vector of intercepts (11) yields the difference of these values $\bar{\mathbf{x}}'$ and their predictions $\bar{\mathbf{x}}'_{\text{pred}}$ by the models (9) of mutual correlation structure

$$\mathbf{A}\bar{\mathbf{x}}' - \mathbf{A}\bar{\mathbf{x}} = \bar{\mathbf{x}}' - \bar{\mathbf{x}}'_{\text{pred}}$$

From this equation, the vector of predictions adjusted by the correlation structure is presented as follows:

$$\bar{\mathbf{x}}'_{\text{pred}} = \left(\mathbf{I} - \mathbf{A}\right)\bar{\mathbf{x}}' + \mathbf{A}\bar{\mathbf{x}} \qquad (12)$$

where $\mathbf{I}$ is the nth order identity matrix. Then vector of updated relative changes equals

$$\mathbf{f}' = \left(\bar{\mathbf{x}}'_{\text{pred}} / \bar{\mathbf{x}}\right) - 1 \qquad (13)$$

and can be used in (7) for estimating the relative change in the DV.

System of Differential Equations for Adjusting Independent Variables

Another way of performing predictor mutual adjustment consists in modeling with a system of equations defining change in each predictor as an aggregate of the values of all the predictors:

$$\begin{cases} \dfrac{dx_1}{dz} = b_{10} + b_{11}x_1 + b_{12}x_2 + \ldots + b_{1n}x_n \\[2mm] \dfrac{dx_2}{dz} = b_{20} + b_{21}x_1 + b_{22}x_2 + \ldots + b_{2n}x_n \\[2mm] \quad\vdots \\[2mm] \dfrac{dx_n}{dz} = b_{n0} + b_{n1}x_1 + b_{n2}x_2 + \ldots + b_{nn}x_n \end{cases} \tag{14}$$

Each predictor derivative on the left-hand side of (14) is presented as a linear model by all the variables on the right-hand side. The derivatives are taken by some variable z identifying a general direction in which all predictors vary. For instance, if observations are gathered in time, it can serve as this profiling variable; price can be another example. Otherwise, if there is no evident variable which can be used for trending other variables by it, the principal component analysis (PCA) approach can be applied to the predictors:

$$\mathbf{C}_{xx}\boldsymbol{\alpha} = \mu\boldsymbol{\alpha} \tag{15}$$

where μ and $\boldsymbol{\alpha}$ are the eigenvalues and eigenvectors of the covariance matrix $\mathbf{C}_{xx}$ (15). Scores by the first component $\boldsymbol{\alpha}_1$ with maximum variance μ_1 can be taken as the variable z defining direction of the main variability of all the xs combined:

$$\mathbf{z} = \mathbf{X}\boldsymbol{\alpha}_1 \tag{16}$$

Then in practical terms, construct a smoothed nonlinear trend of each $\mathbf{x}$ by the new variable $\mathbf{z}$ (for instance, using the "loess" function available in the R software), order all the observation points by ascending values of $\mathbf{z}$, and find the derivatives on the left-hand side of (14) as change in each two subsequent values of each

predictor $\mathbf{x}_j$ by increment in $\mathbf{z}$ values. In the next step, find coefficients in each equation in (14) separately as a linear regression. Having all the coefficients b_{jk} in (14), consider simultaneously solving of this system of linear differential equations.

Similar systems of homogeneous (without intercepts) equations are used in statistics for description of discrete state and continuous time Markov models presented as Chapman-Kolmogorov differential equations describing stochastic processes. As it is well known (e.g., Bellman, 1960; Pennisi, 1972), the solution of a homogeneous linear system of differential equations with constant coefficients can be presented as follows:

$$\mathbf{x}(\mathbf{z}) = \mathbf{P}\operatorname{diag}\left(\exp\left(\lambda_j \mathbf{z}\right)\right)\mathbf{c} \qquad (17)$$

where $\mathbf{c}$ is a vector of constants, and λ_j are the eigenvalues and $\mathbf{P}$ is the corresponding matrix of columns $\mathbf{p}_j$ of eigenvectors obtained in solving the problem:

$$\mathbf{B}\mathbf{p} = \lambda\mathbf{p} \qquad (18)$$

The eigenproblem (18) is considered for the matrix $\mathbf{B}$ of the coefficients b_{jk} (without intercepts b_{j0}) at the right-hand side of (14). For the value $\mathbf{z} = 0$ (corresponding to the mean of PCA scores) the solution (17) reduces to the vector of initial conditions $\mathbf{x}(0) = \mathbf{Pc}$, so we obtain the vector of the constants $\mathbf{c} = \mathbf{P}^{-1}\mathbf{x}(0)$. Using it in (17), the general solution of a homogeneous system of differential equations can be represented as

$$\mathbf{x}(\mathbf{z}) = \mathbf{P}\operatorname{diag}\left(\exp\left(\lambda_j \mathbf{z}\right)\right)\mathbf{P}^{-1}\mathbf{x}(0) \qquad (19)$$

The expression $\mathbf{P}\operatorname{diag}(\exp(\lambda_j\mathbf{z}))\mathbf{P}^{-1}\mathbf{x}(0)$ in (19) is known as matrix exponent. Each component of the vector $\mathbf{x}(\mathbf{z})$ is a linear combination of the exponents $\exp(\lambda_j\mathbf{z})$ which behave in accordance with the specific values of λ_j in the eigenproblem (18). For a real matrix $\mathbf{B}$ the eigenvalues (18) are real numbers or conjugated pairs of complex numbers, which correspond to exponents and to oscillating sine and cosine parts of functions. There also can be polynomial items corresponded to equal eigenvalues, although in practical numerical evaluations such cases are rare. The eigenvectors $\mathbf{p}$ corresponded to the complex eigenvalues are also complex, but the total expression (19) yields real values.

When a homogeneous system of equations is solved in (19), the system (14) with intercepts can be solved as well. Besides exponents, a solution of a nonhomogeneous system can contain additional polynomial by $\mathbf{z}$ items (see the handbook by Kamke (1959), an example eq. 8.10). Taking exponents with already-known parameters λ_j with a polynomial part as a theoretical model of each predictor dependence on the aggregate $\mathbf{z}$, we construct the regressions $\mathbf{x}_j(\mathbf{z})$. For a set of given values $\bar{\mathbf{x}}'$ for predictors (5), we find the corresponding values of PCA scores $\mathbf{z}$ (16) and use them in the models $\mathbf{x}_j(\mathbf{z})$ to adjust predictors by their mutual structure:

$$\bar{\mathbf{x}}'_{j,\text{pred}} = \mathbf{x}_j(\mathbf{z}) = \mathbf{x}_j(\bar{\mathbf{x}}'\boldsymbol{\alpha}_1) \tag{20}$$

The adjusted values (20) are used in estimation of the vector of updated relative changes (13) and then in (7) for finding the relative change in the DV.

Numerical Example

For a numerical example, a data set of 242 respondents from a real marketing research project on customer satisfaction with a service center is considered (this data was used in Lipovetsky and Conklin (2001) for SVR modeling). The variables are measured in a Likert 7-point scale and they are: y – overall satisfaction of clients with the company; $x1$ – customer satisfaction with service representatives; $x2$ – service representatives are courteous; $x3$ – they provide all the needed information; $x4$ – they give quick response; $x5$ – they show care with customer problems; $x6$ – they are accurate in the answers; $x7$ – they take all the necessary actions.

Table 1. Correlations and regressions

Variable	Mean	cor(y, x)	Coefficients	OLS	SVR
y	5.8554	1.0000	$a0$	1.5584	1.1359
$x1$	6.3223	0.5432	$a1$	0.2831	0.1305
$x2$	6.5909	0.4503	$a2$	-0.0292	0.1013
$x3$	6.2727	0.5451	$a3$	0.1792	0.1126
$x4$	6.4339	0.4335	$a4$	-0.0508	0.0955
$x5$	6.1777	0.5110	$a5$	0.0506	0.0892
$x6$	6.3843	0.5462	$a6$	0.2226	0.1292
$x7$	6.2562	0.5045	$a7$	0.0270	0.0849
			R^2	0.3560	0.3447

Table 2. Correlations among the predictors

	x1	x2	x3	x4	x5	x6	x7
x1	1.0000	0.7485	0.6981	0.5891	0.8274	0.7058	0.8023
x2	0.7485	1.0000	0.6277	0.5275	0.7283	0.6853	0.6587
x3	0.6981	0.6277	1.0000	0.7691	0.7156	0.8952	0.7600
x4	0.5891	0.5275	0.7691	1.0000	0.6225	0.7592	0.5912
x5	0.8274	0.7283	0.7156	0.6225	1.0000	0.7161	0.7527
x6	0.7058	0.6853	0.8952	0.7592	0.7161	1.0000	0.6915
x7	0.8023	0.6587	0.7600	0.5912	0.7527	0.6915	1.0000

Table 3. Change in DV by the given and adjusted changes in predictors, in %

Variable	Given increase	Adjusted increase calculated by		
		OLS	SVR	Diff
x1	5.000	5.509	4.219	6.045
x2	4.000	3.195	2.788	4.142
x3	2.000	2.181	4.281	0.517
x4	3.000	1.614	2.450	5.716
x5	10.000	4.802	3.980	0.570
x6	1.000	2.799	3.971	5.068
x7	4.000	5.046	4.720	8.373
y_{pred} OLS	2.505	2.989		2.993
y_{pred} SVR	3.161		3.047	3.513

Presented in Table 1 are mean values of the variables, the pair correlations of y with xs, and coefficients of OLS and SVR models with their coefficients of multiple determination R^2. The DV is correlated with all IVs rather evenly, so each variable could be useful in customer satisfaction impact. In the OLS model some predictor parameters are very close to zero and two of them are negative, but in SVR all parameters are more evenly distributed and all positive. In contrast to OLS, the SVR results are meaningful by all individual coefficients of regression. Judging by the mean values, it could be possible to change the overall satisfaction by improving the predictors.

Correlations between the predictors are shown in Table 2, and they are rather high, so the structure of correlated changes should be taken into account in prediction.

Suppose managers elaborate an improvement program that can result in increasing the mean values in each of the seven predictors by 5, 4, 2, 3, 10, 1, and 4 percent, respectively. The natural question is – what percent of change can be reached for the overall satisfaction by taking these measures? Table 3 in the first

numerical column presents these given predictors' increase, and the corresponding results y_{pred} of the DV change estimated as 2.51% and 3.16% by the OLS and SVR models, respectively, are in the two bottom lines.

Shown in the last three columns in Table 3 are the predictor changes adjusted by the correlation structure using OLS, SVR, and differential equations (denoted Diff in the table). With the predictor increments found in OLS and SVR adjustment by correlation structure, the change in DV equals 2.99% and 3.05%, respectively. Using predictors from the last column of Diff adjustment, we can estimate by both OLS and SVR the yield in DV as 2.99% and 3.51%, respectively. In general, the adjustment increase the output prediction y_{pred}, especially in Diff estimations. To accept a more conservative expectation of the yield among all those adjusted by correlation structure, we can take the SVR prediction y_{pred} of 3.05%. The mean increase is then $5.8554(1 + 0.0305) = 6.034$, which is already above the next level of the seven-point Likert scale of overall satisfaction of clients with company.

Summary

A problem of estimation of a percent change in the dependent outcome variable due to changes in predictors, especially when those are correlated, was considered. Several questions were studied, including formulation of the problem via relation between mean values of all variables, and adjustment of predictors by their correlation structure in the ordinary least squares regression, Shapley value regression, and a model based on solving a system of differential equations. Numerical estimations performed for a real marketing research data set demonstrate meaningful results. Future research can include estimation of the percent change of the outcome due to changes in only a subset of predictors which can be reset almost precisely to some new values (control variables) whereas others cannot be manipulated; however, these would change anyway as a result of their natural correlations with control variables. Lift in the binary outcome and measures improvement in some utility due to selecting observations based on predicted performance (e.g. percent of top performers as predicted by logistic regression) versus a percent change by random selection can be studied as well. In the context of linear regression it is also possible to select, say, the 10% of units with the highest predicted outcome and compare with that 10% of units selected at random, then select the top 20% vs. another randomly selected 20%, etc, which would allow constructing a lift curve. The considered approach can be very useful in practical applications required estimation of the percent change of dependent variable by the

change in predictors in various problems of applied statistical modeling and prediction.

References

Bellman, R. (1960). *Introduction to matrix analysis*. New York, NY: McGraw-Hill.

Gibson, W. A. (1962). Orthogonal predictors: A possible resolution of the Hoffman-Ward controversy. *Psychological Reports, 11*(1), 32-34. doi: 10.2466/pr0.1962.11.1.32

Grapentine, A. (1997). Managing multicollinearity. *Marketing Research, 9*(3), 11-21.

Hoerl, A. E., & Kennard, R. W. (1970). Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics, 12*(1), 55-67. doi: 10.2307/1267351

Johnson, R. M. (1966). The minimal transformation to orthonormality. *Psychometrika, 31*(1), 61-66. doi: 10.1007/bf02289457

Johnson, J. W. (2000). A heuristic method for estimating the relative weight of predictor variables in multiple regression. *Multivariate Behavioral Research, 35*(1), 1-19. doi: 10.1207/s15327906mbr3501_1

Kamke, E. (1959). *Differentialgleichungen losungsmethoden und losungen, I: Gewöhnliche differentialgleichungen*. Leipzig, Germany: B. G. Teubner.

Kendall, M. G., & Stuart, A. (1973). *The advanced theory of statistics* (Vol. 2). New York, NY: Hafner Publishing.

Lipovetsky, S. (2010). Enhanced ridge regressions. *Mathematical and Computer Modelling, 51*(5-6), 338-348. doi: 10.1016/j.mcm.2009.12.028

Lipovetsky, S. (2013). How good is best? Multivariate case of Ehrenberg-Weisberg analysis of residual errors in competing regressions. *Journal of Modern Applied Statistical Methods, 12*(2), 242-255. doi: 10.22237/jmasm/1383279180

Lipovetsky, S., & Conklin, M. (2001). Analysis of regression in game theory approach. *Applied Stochastic Models in Business and Industry, 17*(4), 319-330. doi: 10.1002/asmb.446

Lipovetsky, S., & Conklin, M. (2004). Enhance-synergism and suppression effects in multiple regression. *International Journal of Mathematical Education in Science and Technology, 35*(3), 391-402. doi: 10.1080/00207390310001658492

Lipovetsky, S., & Conklin, M. (2014). Predictor relative importance and matching regression parameters. *Journal of Applied Statistics, 42*(5), 1017-1031. doi: 10.1080/02664763.2014.994480

Mason, C. H., & Perreault, W. D. (1991). Collinearity, power, and interpretation of multiple regression analysis. *Journal of Marketing Research, 28*(3), 268-280. doi: 10.2307/3172863

Pennisi, L. L. (1972). *Elements of ordinary differential equations*. New York, NY: Holt.

Roth, A. E. (Ed.). (1988). *The Shapley value: Essays in honor of Lloyd S. Shapley*. Cambridge, UK: Cambridge University Press. doi: 10.1017/cbo9780511528446

Shapley, L. S. (1953). A value for *n*-person games. In H. W. Kuhn & A. W. Tucker (Eds.). *Contribution to the theory of games (AM-28) II* (pp. 307-318). Princeton, NJ: Princeton University Press. doi: 10.1515/9781400881970-018