A Diagrammatic Representation of Multiple Regression Model Diagnostics for Teaching and Research

Mahayaudin M. Mansor

School of Mathematical Sciences
College of Computing
Informatics and Mathematics
Universiti Teknologi MARA
40450 Shah Alam, Malaysia

Email: maha@uitm.edu.my

Abstract

This contribution features a flowchart representing a general procedure for modeling a response variable that is influenced by several explanatory variables. By modeling, the response values can be estimated based on the values of a set of explanatory variables referenced from literature or have priori with the response variable. However, multivariate regression is an intricate statistical procedure and can be intimidating to many, particularly students and some researchers from a non-statistics background. In addition, advances in data collection methods and digital information facilitate the availability of more exploratory variables, thus making the functional forms of regression more complex. Hence, the objective of the proposed flowchart is to visualize the multivariate regression procedure in a sequence of consolidated actions and processes focusing on diagnostics for multicollinearity, goodness-of-fit, and the underlying assumptions of normality, independence, homoscedasticity, and non-autocorrelation for the residuals. The other advantage is that the flowchart emphasizes not only quantitative but qualitative assessments prior to the establishment of a well-specified model. Consequently, the observed and estimated responses are plotted in a single plot, as many are unable to comprehend the discrepancies with the goodness-of-fit and error measures visually. Whether the response variable is best represented as either a linear or non-linear process, a well-specified model is desired to estimate values that match or are closer to the observed responses, producing more accurate outcomes. This flowchart covers the most important multivariate modeling steps and assumptions and, therefore, can be used as guided notes or a part of instructor-prepared handouts to promote active engagement during lectures, for independent study, or while programming related research. To my knowledge, a similar flowchart has yet to be produced. Besides patenting, there is a possibility to develop a fully automated R package for this work.

Keywords: Qualitative assessment, statistical test, multicollinearity, stepwise procedure, flowchart.

TEACHING AND LEARNING POSTER IDEAS
TALPI 2023
"TOWARDS INNOVATIVE GLOBAL TRENDS IN EDUCATION"
VIRTUAL COMPETITION

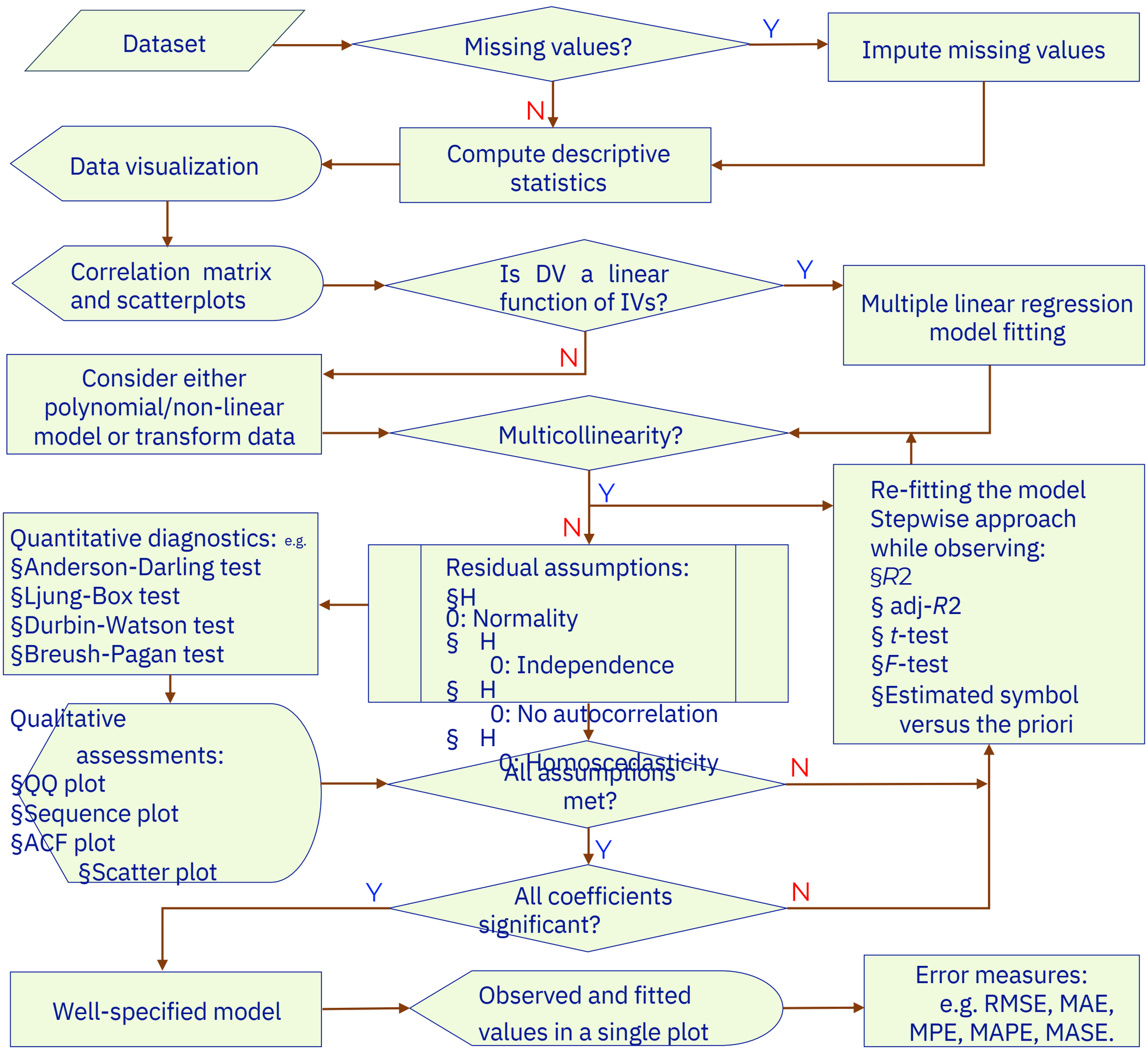
A DIAGRAMMATIC REPRESENTATION OF MULTIPLE REGRESSION MODEL DIAGNOSTICS FOR TEACHING AND RESEARCH

Mahayaudin M. Mansor[1]

[1]School of Mathematical Sciences, College of Computing, Informatics and Mathematics, Universiti Teknologi MARA, 40450 Shah Alam, Malaysia
Corresponding email: maha@uitm.edu.my

This contribution features a flowchart representing a general procedure for modelling a response variable that is influenced by several explanatory variables. By modelling, the response values can be estimated based on the values of a set of explanatory variables referenced from literature or have priori with the response variable. However, multivariate regression is an intricate statistical procedure and can be intimidating to many, particularly students and some researchers from a non-statistics background. In addition, advances in data collection methods and digital information facilitate the availability of more exploratory variables thus making the functional forms of regression more complex. Hence, the objective of the proposed flowchart is to visualize the multivariate regression procedure in a sequence of consolidated actions and processes focusing on diagnostics for multicollinearity, goodness-of-fit, and the underlying assumptions of normality, independence, homoscedasticity, and non-autocorrelation for the residuals. The other advantage is that the flowchart emphasizes not only quantitative but qualitative assessments prior to the establishment of a well-specified model. Consequently, the observed and estimated responses are plotted in a single plot as many are unable to visually comprehend the discrepancies with the goodness-of-fit and error measures. Regardless of whether the response variable is best represented as either a linear or non-linear process, a well-specified model is desired to estimate values that match or are closer to the observed responses, hence producing more accurate outcomes. This flowchart covers the most important multivariate modelling steps and assumptions, therefore can be used as guided notes or a part of instructor-prepared handouts to promote active engagement during lectures, for independent study, or while programming related research.To the best of my knowledge, there is no similar flowchart that has been produced. Besides patenting, there is a possibility to develop a fully automated *R* package for this work.

Keywords: Qualitative assessment, statistical test, multicollinearity, stepwise procedure, flowchart.



Pejabat Perpustakaan
Library Office

Universiti Teknologi MARA
Cawangan Perak
Kampus Seri Iskandar
32610 Bandar Baru Seri Iskandar,
Perak Darul Ridzuan, MALAYSIA
Tel: (+605) 374 2093/2453 Faks: (+605) 374 2299

UNIVERSITI TEKNOLOGI MARA

Surat kami : 700-KPK (PRP.UP.1/20/1)
Tarikh : 20 Januari 2023

Prof. Madya Dr. Nur Hisham Ibrahim
Rektor
Universiti Teknologi MARA
Cawangan Perak

Tuan,

PERMOHONAN KELULUSAN MEMUAT NAIK PENERBITAN UiTM CAWANGAN PERAK MELALUI REPOSITORI INSTITUSI UiTM (IR)

Perkara di atas adalah dirujuk.

2. Adalah dimaklumkan bahawa pihak kami ingin memohon kelulusan tuan untuk mengimbas (*digitize*) dan memuat naik semua jenis penerbitan di bawah UiTM Cawangan Perak melalui Repositori Institusi UiTM, PTAR.

3. Tujuan permohonan ini adalah bagi membolehkan akses yang lebih meluas oleh pengguna perpustakaan terhadap semua maklumat yang terkandung di dalam penerbitan melalui laman Web PTAR UiTM Cawangan Perak.

Kelulusan daripada pihak tuan dalam perkara ini amat dihargai.

Sekian, terima kasih.

"BERKHIDMAT UNTUK NEGARA"

Setuju.

27.1.2023

PROF. MADYA DR. NUR HISHAM IBRAHIM
REKTOR
UNIVERSITI TEKNOLOGI MARA
CAWANGAN PERAK
KAMPUS SERI ISKANDAR

Saya yang menjalankan amanah,

SITI BASRIYAH SHAIK BAHARUDIN
Timbalan Ketua Pustakawan

nar