Journal de Ciencia e Ingeniería, vol. 12, no. 1, pp. 13-19, 2020
http://jci.uniautonoma.edu.co
DOI: https://doi.org/10.46571/JCI.2020.1.2

Corporación Universitaria Autónoma del Cauca
ISSN 2145-2628, e-ISSN 2539-066X

Lagrange Multiplier Tests in Applied Research

Test de multiplicadores de Lagrange en investigación aplicada

José Gabriel Astaiza-Gómez[1]
Departamento de Ciencias Económicas, Universidad del Cauca, Colombia

Abstract. Applied research requires the usage of the proper statistics for hypothesis testing. Constrained optimization problems provide a framework that enables the researcher to build a statistic that fits his data and hypothesis at hand. In this paper I show some of the necessary conditions to obtain a Lagrange Multiplier test as well as some popular applications in order to highlight the usefulness of the test when the researcher must rely in asymptotic theory and to help the reader in the construction of a test in applied work.

Keywords. Likelihood function; Fisher information matrix; regularity conditions.

Resumen. La investigacioń aplicada requiere la utilización de los estadísticos apropiados para probar hipótesis. Los problemas de optimización restringida brindan un marco que le permite al investigador construir un estadístico que tenga en cuenta la naturaleza de sus datos y la hipótesis que se desea probar. En este artículo muestro algunas de las condiciones necesarias para obtener un test de Multiplicadores de Lagrange así como también algunas aplicaciones populares, con el propósito de resaltar la utilidad del test cuando el investigador debe soportarse en teoría asintótica y ayudar al lector en la construcción de un test en investigación aplicada.

Palabras Claves. Funcioń de verosimilitud; matriz de información de Fisher; condiciones de regularidad.

How to cite. R. J. G. Astaiza-Gómez, "Lagrange Multiplier Tests in Applied Research", *Jou. Cie. Ing.*, vol. 12, no. 1, pp. 13-19, 2020. doi:10.46571/JCI.2020.1.2

Received: 28/09/2019 **Revised:** 05/03/2020 **Accepted:** 22/03/2020

1. Introduction

Hypothesis testing, or the use of statistics in order to reject hypotheses, is the main instrument for empirical scientific research. For the researcher, who must find the statistic that fits the hypothesis at hand and the nature of the data, using a Lagrangean expression and the corresponding Lagrange multipliers in order to build the proper statistic from asymptotic theory, has shown to be useful for several situations.

The Lagrange (LM) tests are build upon the distribution of stochastic Lagrange multipliers, obtained from the solution of maximizing the likelihood function in a constrained optimization problem and are asymptotically equivalent to Wald and Likelihood Ratio (LR) tests[2]. These

[1] *e-mail: jgastaiza@unicauca.edu.co*

[2] Interestingly, the LM statistic is always less or equal to the LR statistic, which in turn is always less or equal to the Wald statistic.

three tests are the asymptotically locally most powerful invariant tests, although the LM test has some adventages over the other two: it is the least expensive computationally, its exact distribution for small samples can sometimes be found and its computation requires no more than the residuals from a least squares regression.

Among the popular tests used in applied work that belong to the LM family we can find the Godfrey test for autocorrelation, the Jarque-Bera normality test, the χ^2 contingency table test and the Wu-Hausman test for unobserved variable. Given the popularity and advantages of the LM procedure to build statistics for hypothesis testing, in this paper I describe some required conditions for the statistic to exist as well as the procedure for its application in popular cases. Thus, this paper may help the reader in the construction of a test in applied work.

This document is divided in five sections including the introduction. In section two, I review the literature on test statistics based on asymptotic theory. In section three I describe some of the fundamental assumptions required to develop the LM statistic. In section four I show some of the popular applications that rely on the LM test for hypothesis testing and in section five I conclude.

2. Related Literature

By using the framework of Lagrange multipliers to model a constrained optimization problem, [1] proposed a test based on the Lagrange multipliers obtained from finding the parameters that maximize the likelihood function, today known as the LM test. Until then, the problem of identifying whether the parameter values of a distribution belonged to a subset of parameters, under the assupmtion of independence, was solved mainly through the Neyman - Pearson Likelihood Ratio test [13]. Becasuse of its adventages, the LM test, which was separately formulated by [12], has gained popularity with time and several important applications have been developed since its appearance.

Under the regularity conditions of Maximum Likelihood, the LM test is asymptotically equivalent to the Wald [16] and LR tests, and the three of them share the property of being the asymptotically locally most powerful invariant tests [5]. Nevertheless, the LM test has the the least computational costs. In addition, its exact distribution for small samples can be obtained in specific cases and having the residuals from a simple least squares regression is enough for its estimation. As a consequence, many popular tests have been developed using the LM framework, such as "the Wu-Hausman unobserved variable test and Kmenta's test for a Cobb Douglas production function (see[6]), the Chow test, Andrews' functional form test [2] , the χ^2 contingency table test and the partial autocorrelation function"[4].

3. The Lagrange Multiplier Test

Let $\boldsymbol{x} = (x_1, x_2, \dots, x_n)$ be a set of n independent observations on a random variable with distribution function F, which in turn depends on k parameters $\theta_1, \theta_2, \dots, \theta_k$, and let $\boldsymbol{\theta} = (\theta_1^0, \theta_2^0, \dots, \theta_k^0)$ be the vector of true, but unknown, parameter values. Also, there is a set of restrictions $h_j(\theta_1, \theta_2, \dots, \theta_k) = 0$, $j = 1, 2, \dots, r$, $r < k$, so that the Lagrange theorem can be applied. Define the log-likelihood function as

$$L_n = \sum_{i=1}^{n} log f(x_i, \boldsymbol{\theta}). \tag{1}$$

where $f(\cdot, \cdot)$ is the density function. By finding arguments that maximize the log-likelihood function, the Lagrange statistic can be expressed as (see [13] and [4])

$$\xi^{LM} = \frac{1}{n} \hat{\lambda}' g(\hat{\theta})' I(\hat{\theta})^{-1} \hat{\lambda} g(\hat{\theta}) \sim \chi_r^2 \tag{2}$$

where the hats indicate the solution values, $\hat{\lambda}$ is the vector of Lagrange multipliers that solve the problem, $g(\hat{\theta})$ is the Jacobian matrix of the restraints and $I(\hat{\boldsymbol{\theta}})^{-1}$ is the inverse of the Fisher information matrix.

The above expression can also be expressed as (see e.g. [15])

$$\xi^{LM} = \mathbf{1}'S(S'S)^{-1}S'\mathbf{1} \tag{3}$$

where $\mathbf{1}$ is a vector of ones of size $(n \times 1)$ and S is a $(n \times k)$ matrix whose element $(1,1)$ corresponds to the derivative, with respect to the first parameter, of the log of the density function in the first observation. Similarly, the element in the position (n,k) is the derivative of the log of the density function in the observation n, with respect to the $k-th$ parameter. In addition, each element is evaluated at the null hypothesis.

Importantly, for the maximization problem to have a unique solution, several conditions must hold. Among them, it is assumed that the set of parameters is a convex compact subset, that the log of the density function and the constraint functions are continuous on the set of parameters and that $\dfrac{\partial log\ f(\cdot,\theta)}{\partial \theta_i}$ $(i = 1, 2, ..., k)$ exist for all θ. More generally, for the construction of the LM statistic it must be that 1) the likelihood function satisfies standard regularity conditions so that it is possible to interchange the derivative and the integral and make Taylor expansions, 2) the Fisher information matrix is not singular so that the parameters are locally identified, 3) the Lagrange conditions hold, and 4)the Central Limit Theorem can be applied to the scores.

4. Applications

In this section I show three popular applications of the Lagrange Multiplier framework. To be more concrete, I revise the tests for hypotheses on the parameters on independent variables and tests for characteristics on the unobservables.

4.1. Testing Restrictions on Conditional Means

Consider the following model:

$$y_i = x_{1,i}\theta_1 + x_{2,i}\theta_2 + \epsilon_i \tag{4}$$

for $i = 1, 2, 3..., n$, $n \in \mathbb{N}$, and $\epsilon_i \sim n.i.i.d(0, \sigma^2)$. Also, $y_i, x_{1,i}$ and $x_{2,i}$ are observable variables with $x_{1,i}, x_{2,i}$ exogenous, and θ_1, θ_2 are the population parameters. Notice that the statistical model 4 fits in the framework of the LM test, i.e. the solution of the optimization problem exists and is unique. The log-likelihood function is

$$L_n = \sum_{i=1}^{n} logf(y_i, \boldsymbol{\theta}) = -\frac{n}{2}log2\pi - \frac{n}{2}log\sigma^2 - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i - x_{1,i}\theta_1 - x_{2,i}\theta_2)^2 \tag{5}$$

where $\boldsymbol{\theta}$ is the vector of parameters and $f(\cdot,\cdot)$ is the density of a normal distribution with mean $x_{1,i}\theta_1 + x_{2,i}\theta_2$ and variance σ^2. Since $\frac{n}{2}log2\pi$ is a constant, maximizing 5 is equivalent to maximize

$$-\frac{n}{2}log\sigma^2 - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i - x_{1,i}\theta_1 - x_{2,i}\theta_2)^2 \tag{6}$$

Under the null hypothesis that $\theta_2 = 0$, the Lagrange function and first order conditions are:

$$\begin{array}{rcl} L = & L_n - \lambda\theta_2 & \\ \sum_{i=1}^{n} \dfrac{\partial log f(\cdot,\cdot)}{\partial\theta_1} = & \sum_{i=1}^{n} \dfrac{\epsilon_i}{\sigma^2} x_{1,i} = & 0 \\ \sum_{i=1}^{n} \dfrac{\partial log f(\cdot,\cdot)}{\partial\theta_2} = & \sum_{i=1}^{n} \dfrac{\epsilon_i}{\sigma^2} x_{2,i} - \lambda = & 0 \\ \sum_{i=1}^{n} \dfrac{\partial log f(\cdot,\cdot)}{\partial\sigma^2} = & \sum_{i=1}^{n} \epsilon_i^2 - \dfrac{n}{2\sigma^2} = & 0 \end{array} \tag{7}$$

Solving for λ, the stochastic Lagrange multiplier is

$$\hat{\lambda} = \sum_{i=1}^{n} \frac{\partial log f(\cdot,\cdot)}{\partial\theta_2} \tag{8}$$

and the LM statistic ξ^{LM} is defined as in equation 3, with $\hat{\epsilon}_i = y_i - x_{1,i}\hat{\theta}_1$ and

$$S = \begin{bmatrix} \hat{\epsilon}_1 x_{1,1} & \hat{\epsilon}_1 x_{2,1} \\ \vdots & \vdots \\ \hat{\epsilon}_n x_{1,n} & \hat{\epsilon}_n x_{2,n} \end{bmatrix} \tag{9}$$

4.2. Testing for Constant Variance

Now consider the following model:

$$y_t = \boldsymbol{x}_t'\boldsymbol{\theta} + \epsilon_t \tag{10}$$

where the vector $\boldsymbol{x}_t'$ is exogenous for $t = 1, 2, 3..., n$, $n \in \mathbb{N}$. Different to the previous section, we will not assume that ϵ_t has a constant variance but rather $\epsilon_t \sim n.i.i.d(0, \sigma_t^2)$ where σ_t^2 is a function of some exogenous variables $\boldsymbol{z}_t$ and parameters $\boldsymbol{\alpha}$ not related to $\boldsymbol{\theta}$:

$$\sigma_t^2 = h(\boldsymbol{z}_t'\boldsymbol{\alpha}) \tag{11}$$

where $h(\cdot)$ is twice differenciable, the vector of parameters $\boldsymbol{\alpha}$ is of size $(p \times 1)$ and the first element of the vector $\boldsymbol{z}_t$ is the unit. Ommiting its constants, the log-likelihood function is

$$L_n = -\frac{n}{2} log 2\pi - \sum_{t=1}^{n} log\sigma_t^2 - \frac{1}{2}\sum_{i=1}^{n} \frac{1}{\sigma_t^2}(y_t - \boldsymbol{x}_t'\boldsymbol{\theta})^2 \tag{12}$$

We would like to test the null hypothesis of constant variance, which is equivalent to test for $\alpha_2 = \alpha_3 = \ldots = \alpha_p = 0$, i.e.:

$$\sigma_t^2 = \sigma^2 = h(\alpha_1) \tag{13}$$

Following the notation in [12] , the LM statistic is

$$\xi^{LM} = \left[\frac{\partial\hat{L}_n}{\partial\boldsymbol{\alpha}}\right]' \hat{I}^{-1} \left[\frac{\partial\hat{L}_n}{\partial\boldsymbol{\alpha}}\right] \tag{14}$$

where the hat indicates that the function is being evaluated at the constrained parameters and I is the Fisher information matrix

$$I = -\mathbb{E}\left[\frac{\partial^2 L_n}{\partial\boldsymbol{\alpha}\partial\boldsymbol{\alpha}'}\right] \tag{15}$$

This test is the homoskedasticity test of [3] .

4.3. Testing for Autocorrelation

Another popular application of the LM framework is to test for correlation of the error terms. More specifically, the test for autocorrelation of [9] , tests the null hypotesis of independent residuals versus the alternative hypothesis that the error terms follow an autorregresive process. Consider again the model of the previous section:

$$y_t = \boldsymbol{x}_t'\boldsymbol{\theta} + \epsilon_t \tag{16}$$

where the vector $\boldsymbol{x}_t'$ is exogenous for $t = 1, 2, 3..., n$, $n \in \mathbb{N}$. Under the alternative hypothesis of autocorrelation:

$$u_t = \epsilon_t + \rho_1\epsilon_{t-1} + \rho_2\epsilon_{t-2} + \ldots + \rho_p\epsilon_{t-p} \tag{17}$$

where $u_t \sim n.i.i.d(0, \sigma_u^2)$. We can see that the null hypothesis of $\epsilon_t \sim n.i.i.d.(0, \sigma^2)$ implies that $\rho_1 = \ldots = \rho_p = 0$. For ease of exposition, define

$$M = \begin{bmatrix} 1 & & & & 0 \\ \rho_1 & 1 & & & \\ \vdots & \ddots & \ddots & & \\ \rho_p & \ldots & \ddots & 1 & \\ 0 & \rho_p & \ldots & \rho_1 & 1 \end{bmatrix}$$

and $\boldsymbol{u} = M\boldsymbol{\epsilon}$. The log-likelihood function is

$$L_n = -\frac{1}{2}log\sigma_u^2 - \frac{1}{2n\sigma^2}\boldsymbol{u}'\boldsymbol{u}$$

and the concentrated function with respect to σ_u^2 is

$$L_n^c = -\frac{-1}{2}log\frac{\boldsymbol{u}'\boldsymbol{u}}{n}$$

The Lagrange function of the constrained optimization problem under $\rho_1 = \ldots = \rho_p = 0$ is:

$$L = L_n^c - \boldsymbol{\lambda\rho}$$

where $\boldsymbol{\lambda}$ and $\boldsymbol{\rho}$ are the vectors of Lagrange multipliers and distribution parameters respectively. Using the Lagrange multipliers that solve the problem $\hat{\boldsymbol{\lambda}} = \dfrac{\partial \hat{L}_n^c}{\partial \boldsymbol{\alpha}}$, the researcher can estimate the LM statistic to test the hypotheses.

4.4. Recent Advances and Applications with Observational Data

In order to test whether obesogenic environment accentuates the risk of obesity in adults, [14] gathered a sample of up to 120 000 adults from the UK with information on their body mass index (BMI) as the outcome variable and variables related to genetic risks and the obesogenic environment, including Townsend deprivation index (TDI), TV watching, diet and physical activity. Since the authors suspected that the variance in BMI was higher in individuals in the high-risk environment groups, they tested for heteroscedasticity using the Breusch-Pagan test and used robust standard errors to control for heteroscedasticity. After correcting for heteroscedasticity, [14] found statistical evidence in favor of the hypothesis that "the obesogenic environment accentuates the risk of obesity in genetically susceptible adults" and that "relative social deprivation best captures the aspects of the obesogenic environment" from the set of variables they included in their research. It was important to correct for differences in the variance, since otherwise they results would be inflated. Xie [17] study the relationship between hotel financial performance and management responses to online reviews. The authors use data on quarterly hotel revenues and also count with 22,483 management responses to 76,649 online consumer reviews on TripAdvisor over 26 quarters. Since the authors use time-series data, they carry out an LM tests for autocorrelation as in [9], from which they cannot reject the null hypothesis of no first-order serial correlation. In their research, [17] find that "that providing timely responses to online reviews enhances financial performance of hotels".

Related to research on retinal therapies and using the fly eye, [11] investigate the migratory responses of innate collections of retinal progenitor cells (RPCs) upon extracellular substrates. They tested for normality using the Jarque-Bera test [10] in order to establish whether the data for single cells, small clusters and large clusters was different from normally distributed data.

Recent advances in hypothesis testing using Lagrange Multipliers include [8], to test for heteroskedasticity and autocorrelation in spatial models, and [7] who propose a LM test for financial contagion based on a multivariate generalized normal distribution.

5. Conclusions

In this paper I describe some required conditions for the Lagrange Multiplier statistic to exist as well as the procedure for its application in popular cases. From the assumptions needed for the development of the test, we know that it is not suitable for situations in which the likelihood function is not continuous in the parameter support. Nevertheless, in most situations the continuity assumption is intuitive and given the advantages of the LM test, it is useful for researchers dealing with data, allowing them to build a proper test that fits their data structure and research hypothesis. The cases shown in this paper describe how the LM framework can be applied to situations in which the data generating processes are described by different functional forms, which in turn exposes the usefulness of the test when the researcher must rely in asymptotic theory.

References

[1] J. Aitchison and S. D. Silvey. "Maximum Likelihood Estimation of Parameters Subject to Restraints", *Annals of Mathematical Statistics*, 29 (3): 813-828, (1958).

[2] D. F. Andrews. "A Note on the Selection of Data Transformations", *Biometrika*, 58: 249-254, (1971).

[3] T .S. Breusch and A. R. Pagan. "A Simple Test for Heteroscedasticity and Random Coefficient Variation," *Econometrica*. 47 (5): 1287-1294, (1979).

[4] T .S. Breusch and A. R. Pagan. "The Lagrange multiplier test and its applications to model specification in econometrics," *Rev. Econ. Studies* 97: 239-253, (1980).

[5] R. Engle. "Wald, likelihood ratio, and Lagrange multiplier tests in econometrics", in *Handbook of Econometrics*, vol. 2, Ch. 13 , Z., Griliches and M., D. Intriligator eds., Elsevier: 775-826, (1984).

[6] R. Engle. "The Lagrange multiplier test as a regression diagnostic." Unpublished working paper. University of California, San Diego, (1978).

[7] R. Fry-McKibbin, C. Y. Hsiao, and V. L. Martin. "Joint tests of contagion with applications." *Quantitative Finance* 19(3): 473-490, (2019).

[8] J. Le Gallo, F. A. López, and C. Chasco. "Testing for spatial group-wise heteroskedasticity in spatial autocorrelation regression models: Lagrange multiplier scan tests." *The Annals of Regional Science*: 1-26, (2019).

[9] L. G . Godfrey. "Testing Against General Autoregressive and Moving Average Error Models When the Regressors Include Lagged Dependent Variables", *Econometrica*, 46: 1293-1301, (1978).

[10] C. Jarque and A. Bera. "A Test for Normality of Observations and Regression Residuals", *International Statistical Riview*, 55: 163-172, (1987).

[11] C. D. Pena, et al. "Invertebrate Retinal Progenitors as Regenerative Models in a Microfluidic System." *Cells* 8(10): 1301, (2019).

[12] C. R. Rao. "Large Sample Tests of Statistical Hypotheses Concerning Several Parameters with Application to Problems of Estimation", *Proceedings of the Cambridge Philosophical Society*,44: 50-57, (1948).

[13] S. D. Silvey. "The Lagrangian multiplier test." *The Annals of Mathematical Statistics*, 30, (2): 389-407, (1959).

[14] Tyrrell, Jessica, et al. "Gene-obesogenic environment interactions in the UK Biobank study." *International journal of epidemiology* 46.2: 559-575 (2017).

[15] M. Verbeek. *A Guide to Modern Econometrics*, 2nd edition. Chichester: John Wiley and Sons, (2004).

[16] A. Wald. "Tests of Hypotheses Concerning Several Parameters When the Number of Observations is Large", *Transactions of the American Mathematical Society*, 54: 426-482, (1943).

[17] K. L. Xie, K. K. Fung So, and W. Wang. "Joint effects of management responses and online reviews on hotel financial performance: A data-analytics approach." *International Journal of Hospitality Management*, 62: 101-110 (2017).

Journal de Ciencia e Ingeniería Uniautónoma DEL CAUCA