International Statistical Conference
in Memory of Professor Sik-Yum Lee
Hong Kong, December 2019

S.-Y. Lee's Lagrange Multiplier Test in Structural Modeling: Still Useful?

Peter M. Bentler

University of California, Los Angeles

Sik-Yum Lee: Early Productivity

Bentler, P. M., & Lee, S.-Y. (1975). Some extensions of matrix calculus. *General Systems*, *20*, 145-150.

Bentler, P. M., & Lee, S.-Y. (1978). Matrix derivatives with chain rule and rules for simple, Hadamard, and Kronecker products. *Journal of Mathematical Psychology*, *17*, 255-262.

Bentler, P. M., & Lee, S.-Y. (1978). Statistical aspects of a three-mode factor analysis model. *Psychometrika*, *43*, 343-352.

Bentler, P. M., & Lee, S.-Y. (1979). A statistical development of three-mode factor analysis. *British Journal of Mathematical and Statistical Psychology*, *32*, 87-104.

Bentler, P. M., & Lee, S.-Y. (1979). Newton-Raphson approach to exploratory and confirmatory maximum likelihood factor analysis. *Journal of the Chinese University of Hong Kong*, *5*, 562-573.

Lee, S.-Y., & Jennrich, R. I. (1979). A study of algorithms for covariance structure analysis with specific comparisons using factor analysis. *Psychometrika*, *44*, 99-113.

Lee, S.-Y., & Bentler, P. M. (1980). Functional relations in multidimensional scaling. *British Journal of Mathematical and Statistical Psychology*, *33*, 142-150.

Lee, S.-Y., & Bentler, P. M. (1980). Some asymptotic properties of constrained generalized least squares estimation in covariance structure models. *South African Statistical Journal*, *14*, 121-136.

Outline

- Sik-Yum Lee on Evaluating Restricted Models
- More General Distributions and Model Types
- Early Critiques and Evaluations
- Another Approach: Expected Parameter Change
- Recent Studies and Current Status
- Using a Tiny Bit of Theory
- A Small Simulation
- Conclusion

Sik-Yum Lee on Evaluating Restricted Models

SOME ASYMPTOTIC PROPERTIES OF CONSTRAINED GENERALIZED LEAST SQUARES ESTIMATION IN COVARIANCE STRUCTURE MODELS

Sik-Yum Lee
Department of Mathematics
University Science Centre
The Chinese University of Hong Kong
Hong Kong

P.M. Bentler
Department of Psychology
University of California
Los Angeles
California 90024

Some basic results by Browne (1974) on generalized least squares estimation in the analysis of covariance structures are extended to covariance structures with parameters subject to arbitrary nonlinear constraints. It is shown that the constrained estimators are consistent, asymptotically multivariate normally distributed, and asymptotically equivalent to constrained maximum likelihood estimators. Asymptotic chi-square tests are developed to evaluate appropriate model comparisons. The relationships between the Lagrangian approach and the reparameterization approach are discussed.

$\Sigma_o = \Sigma(\underline{\theta}_o)$ subject to $\underline{h}(\underline{\theta}) = \underline{0}$

$$Q(\underline{\theta}) = 2^{-1} \operatorname{tr} \{(S-\Sigma)V\}^2,$$

$$\dot{Q}(\underline{\tilde{\theta}}) + \tilde{L}'\underline{\tilde{\lambda}} = \underline{0}$$

$$\underline{h}(\underline{\tilde{\theta}}) = \underline{0}, \qquad L = (\partial h_i / \partial \theta_j) \qquad \underline{\tilde{\lambda}}' = (\tilde{\lambda}_1, \ldots, \tilde{\lambda}_r)$$

$$F(\underline{\theta}) = \log|\Sigma| + \operatorname{tr}(S\Sigma^{-1}) - \log|S| - p.$$

$$\dot{F}(\underline{\theta}) + \hat{L}'\underline{\hat{\lambda}} = \underline{0}$$

$$\underline{h}(\underline{\hat{\theta}}) = \underline{0},$$

AITCHISON, J. and SILVEY, D.S. (1958). Maximum likelihood estimation of parameters subject to restraints. *Ann. Math. Statist.*, **29**, 813-828.

Proposition 2

The joint asymptotic distribution of random variables $n^{1/2}(\tilde{\underline{\theta}} - \underline{\theta}_o)$ *and* $n^{1/2}\tilde{\underline{\lambda}}$ *is multivariate normal with zero mean vector and covariance matrix*

$$2\begin{pmatrix} P_o & 0 \\ 0 & -R_o \end{pmatrix},$$

$$P_o = \ddot{M}_o^{-1} - \ddot{M}_o^{-1}L_o'(L_o\ddot{M}_o^{-1}L_o')^{-1}L_o\ddot{M}_o^{-1}$$

$$\ddot{M}_o = M_o + L_o'L_o \qquad M_o = \Delta_o'(\Sigma_o^{-1} \otimes \Sigma_o^{-1})\Delta_o$$

$$R_o = I_r - (L_o\ddot{M}_o^{-1}L_o')^{-1}$$

Proposition 3

The generalized least squares estimator $(\tilde{\underline{\theta}}, \tilde{\underline{\lambda}})$ *is asymptotically equivalent to the maximum likelihood estimator* $(\hat{\underline{\theta}}, \hat{\underline{\lambda}})$.

Proposition 4

The asymptotic distribution of $nQ(\tilde{\underline{\theta}})$ *is chi-square with degrees of freedom* $p(p+1)/2-(q-r)$.

Suppose $\underline{h}(\underline{\theta}) = (\underline{h}^*(\underline{\theta}),\ \underline{h}^{**}(\underline{\theta}))$,

Proposition 5

The asymptotic distribution of $n\cdot\{Q(\tilde{\underline{\theta}}) - Q(\tilde{\underline{\theta}}^*)\}$ *is chi-square with degrees of freedom* $r-j$.

Proposition 6

The asymptotic distribution of $-2^{-1}n\tilde{\underline{\lambda}}'\tilde{R}^{-}\tilde{\underline{\lambda}}$ *under* H_o *is chi-square with degrees of freedom equal to the rank of* R_o, *where*

$$\tilde{R} = I_r - (\tilde{L}\ddot{M}(\tilde{\underline{\theta}})^{-1}\tilde{L}')^{-1} \text{ (see Lemma 3).}$$

It is well known that this test is asymptotically equal to Rao’s (1948) score test. These tests also are asymptotically equal to the Wald and LR chi-square difference tests (see e.g., Buse, 1982). Sik-Yum Lee (1985) developed a Wald test for this situation, while Sörbom (1989) developed the score test (labeled MI or "modification index"; a redefinition of Sörbom's 1975 MI).

The LM test for several omitted parameters can be broken down into a series of 1-df tests. Bentler (1983, 1985) developed a forward stepwise LM procedure where, at each step, the parameter is chosen that will maximally increase the LM chi-square, contingent on those already included. In EQS, this is based on Beaton's (1964) SWEEP operator. Sörbom (1989) also mentioned such an approach.

It seems that the most frequent applications of LM tests in SEM are the following:

- $\theta_i = 0$. Evaluate necessity of an omitted parameter. This is often – maybe almost always – post-hoc.
- $\theta_i - \theta_j = 0$. Evaluate the appropriateness of an equality restriction. This can be a priori.
- In EQS, evaluate constraints across multiple groups such as, for a given parameter, $\theta_i^{(1)} = \theta_i^{(2)} = \ldots = \theta_i^{(g)}$, i.e., differences are zero. This is typically a fully a priori test, e.g., of equal factor loadings across groups.

Simple nonlinear constraints such as $\theta_1 = \theta_2^2$ can be done with phantom variables (Rindskopf, 1984), and do not require constrained optimization. Tang & Bentler

(1998) gave a restricted EM algorithm to compute constrained estimates for structural models with missing data.

More General Distributions and Model Types

Extensions of the LM test to distribution-free and specialized distributions, as well as to misspecified distributions, were given by Bentler & Dijkstra (1985) and Satorra (1989), and implemented in EQS (Bentler, 1985) for general linear constraints. Methodology for nonlinear constraints exists in several SEM programs.

Following Sik-Yum Lee & Tsang (1999) for covariance structures, Bentler, Liang, Tang, & Yuan (2011) developed an EM algorithm for constrained ML estimation for 2-level mean and covariance structure models. Lee's LM test has been extended in EQS to evaluate constraints in this and other model setups.

General nonlinear constraints seem to be infrequently used. More important are nonlinear structural equations, where Sik-Yum Lee and colleagues (e.g., X. Y. Song) have been major contributors.

Lee envisioned the LM test to be used for evaluating a priori hypotheses about parametric constraints. This is the recommended approach in instructional materials and in structural equation programs (e.g., EQS, LISREL, Mplus). As noted, it seems exploratory search and evaluation of $\theta_i = 0$ for a large set of parameters (e.g., factor loadings set to zero, omitted paths, possible error covariances) is the most typical application.

Early Critiques and Evaluations

Cliff (1983) provided an early critique, noting e.g., that ex-post facto analyses are not tests of models. “Long established scientific principles must still be applied.”

MacCallum (1986) showed that recovery of a true model by specification searches is usually difficult, but improved when “when (a) the investigator’s initial model corresponds closely to the true model, (b) the search is allowed to continue even when a statistically plausible model is obtained, (c) the investigator can place valid restrictions on permissible modifications, and (d) a large sample is used.”

Chou & Bentler (1990) reported that "when a correct null hypothesis was embedded in a composite hypothesis which was false, an incremental LM test tended to suggest more parameters than needed to be freed, especially at larger sample sizes. This incorrect behavior of the LM test was correctable by following up the LM test by a W test." This result verified one of MacCallum's conclusions (b).

MacCallum, Roznowski, & Necowitz (1992) evaluated capitalization on chance with MI. "Results demonstrate that over repeated samples, model modifications may be very inconsistent and cross-validation results may behave erratically… (and) lead to skepticism about generalizability of models resulting from data-driven modifications of an initial model.”

Another Approach: Expected Parameter Change

Saris, Satorra, & Sörbom (1987) showed that large MIs (1-df LM test) can be associated with trivial or small misspecifications. They proposed a parameter change statistic to assess size of parameter misspecification, and to use instead of, or in addition to, the MI. Suppose parameter π_i is restricted to π_0 (usually = 0), and $d_{\pi_i} = \partial \ln L(\pi) / \partial \pi_i$ is evaluated at the estimated model. They defined parameter change as

$$\pi_i - \pi_0 = MI / d_{\pi_i} \ ,$$

which, if $\pi_0 = 0$, provides a prediction on the size of $\hat{\pi}_i$ if it were to be included in the model as a free parameter.

Using Lee's terminology, $\hat{\pi}_i = LM_i / n\hat{\lambda}_i$

This index is now called Expected Parameter Change (EPC). A version for multiple simultaneous constraints or fixed parameters was given by Bentler (1989, 1990), Satorra (1989), and Chou & Bentler (1993).

Luijben & Boomsma (1988) showed that the size of an EPC can depend on how variables and latent variables in the models are scaled, making it hard to compare EPCs from various parts of a model.

Kaplan (1989) introduced a standardized EPC (SEPC) and showed in an application "that the MI tends to suggest freeing substantively implausible parameters. The EPC and SEPC, by contrast, suggest freeing substantively interesting parameters."

Chou & Bentler (1993) noted that Kaplan's approach to standardizing EPC did not yield results that are invariant to different scalings of latent and observed variables. The proposed a fully standardized SEPC, which has become accepted. They also proposed a fully standardized version of multivariate EPCs for a set of constraints or fixed parameters.

The above results were obtained decades ago. There have been dozens of related studies since then. But has theory or practice been improved in the meantime?

Recent Studies and Current Status

As far as I can tell, the theory, practical usefulness, and best practices on a priori LM tests, post-hoc model modification, and use of EPC or SEPC are considered to be about the same today as they were in the early years. A few illustrative more recent studies are the following.

Whittacker (2012) studied an 8-variable 2 factor confirmatory factor analysis model. The factor inter-correlation was omitted, and MI and SEPC were used to find an improved model. Simulated conditions varied sample size, factor loading size, and factor inter-correlation size on ability to find the correct model. “The results indicated that, in general, the SEPC outperformed the MI when arriving at the correct confirmatory factor model. However, they performed more similarly as factor loading size, sample size, and misspecified parameter size increased.” Also, joint criteria (significant MI and largest SEPC) “proved to be slightly less accurate than the significant MI,” though “among the set of fixed parameters associated with significant MI values and the largest SEPC value, the correct parameter was known."

In the context of multiple group modeling, Jorgensen (2017) considered the problem of evaluating whether a parameter might be freed simultaneously in all groups. In an empirical study with a 2 group model, he found the 2-df multivariate LM test was effective. Also, a Monte Carlo simulation with a 4-group model "illustrated how (the multivariate LM test) … could limit Type I errors better than traditional 1-df modification indices for individual fixed parameters within each group."

Marcoulides & Falk (2018) made the Tabu heuristic optimization procedure available for model search in R, and illustrated it using a BIC criterion in Lavaan.

Thus:

(1) A priori use of LM tests according to Lee’s theory remains fully justified;
(2) A posteriori use of LM tests in a specification search is still a reasonable option to find possible omitted parameters or alternative models. However, some false positives (true 0’s declared non-zero) seem to be inevitable and validation is needed.
(3) The SEPC tends to be somewhat more accurate than the LM test in identifying misspecified fixed parameters.

Is there any way to improve on current practice?

Using a Tiny Bit of Theory

Although often recommended, using theory to specify which fixed parameters to evaluate by LM test seems hard to do. A simpler, but perhaps feasible, approach is the following:

Evaluate by MI or LM test only those parameters for which an a priori hypothesis can be made on the expected sign (+ or -) of that parameter.

Obviously, this should reduce the number of false positives. It also would provide researchers with a simple way to ignore large but misleading (wrong sign) LM results. Of course, in a few cases a potentially valid insightful, though unexpected, result might then be ignored.

A Small Simulation

CFA model with 21 variables, 3 correlated factors

Simple cluster structure

V1-V7 loading .7 on F1
V8-V14 loading .7 on F2
V15-V21 loading .7 on F3

as well as 3 cross loadings of .3:

V1 on F3; V8 on F1; V15 on F2

F1-F3 intercorrelate .6

Unique variances are such that a correlation matrix results.

The model to be fit excludes the 3 cross-loadings. This is a minimally misspecified model with 42 possible misspecifications. The population Standardized Root Mean square Residual (SRMR) = .039, with the largest standardized residual being .078. (Compare to Marcoulides & Falk, 2018, with SRMR = .09, and 3/11 true/evaluated misspecifications.)

100 samples of size n=100 were drawn from a normally distributed population with population covariance matrix generated by the true model (the 3 cross-loadings of .3 were included). In each sample, the *false* model was estimated by ML, and LM tests and SEPCs were computed. Omitted loadings were hypothesized to be positive in sign.

1. Ordering omitted parameters by size of MI (1-df LM), evaluating the top 5 positive SEPCs:

95% of the top 5 have all 3 true omitted parameters

5% of the top 5 have 2/3 true

2. Ordering omitted parameters by size of MI (1-df LM), evaluating the top 3 positive SEPCs:

64% of the top 3 have all 3 true omitted parameters

36% of the top 3 have 2/3 true

In comparison, the blind forward stepwise LM with a default method (univariate increment n.s.) to stop entering parameters:

74% captured all 3 true; 24% 2/3; and 2% 1/3

I would conclude {LM + directional SEPC} can be an improvement over the stepwise LM test. A larger study seems warranted.

Given that the {LM + directional SEPC} had no automatic procedure to determine the number of parameters to add, and also yielded some false positives, additional further potential improvements could be:

(1) forward stepwise LM entering only parameters with positive SEPCs;

(2) a final backward stepwise Wald test, to remove unnecessary (hopefully, false positive) parameters.

Conclusion

Among his many contributions, Sik-Yum Lee provided the field with a wonderful tool to evaluate restricted structural equation models that has stood the test of time.

Unfortunately, with few exceptions (such as invariance restrictions in multiple group models), substantive theory in fields that use structural models seems often to be poorly developed, making it hard to actually specify many a priori parametric restrictions on models. Thus it is hard to use Lee's results in a statistically correct way.

When a priori models are inadequate, it is often necessary to engage in a specification search for model improvements. Lee's results continue to contribute to this endeavor.

References

Beaton, A. E. (1964). *The use of special matrix operators in statistical calculus*. Educational Testing Service: RB-64-51.

Bentler, P. M. (1983). *Complete SWP with constraints and large W matrix*. UCLA: Technical Note.

Bentler, P. M. (1985). *Theory and implementation of EQS, A structural equations program*. Los Angeles: BMDP Statistical Software.

Bentler, P. M. (1989). *EQS structural equations program manual.* Los Angeles: BMDP Statistical Software.

Bentler, P. M. (1990). Fit indexes, Lagrange multipliers, constraint changes and incomplete data in structural models. *Multivariate Behavioral Research*, *25*, 163-172.

Bentler, P. M., & Dijkstra, T. (1985). Efficient estimation via linearization in structural models. In P. R. Krishnaiah (Ed.), *Multivariate analysis VI* (pp. 9-42). Amsterdam: North-Holland.

Bentler, P. M., Liang, J., Tang, M. –L., & Yuan, K. –H. (2011). Constrained maximum likelihood estimation for two-level mean and covariance structure models. *Educational & Psychological Measurement*, 71, 325–345.

Buse, A. (1982). The likelihood ratio, Wald and Lagrange multiplier tests: An expository note. *American Statistician*, *36*, 153-157.

Chou, C.-P., & Bentler, P. M. (1990). Model modification in covariance structure modeling: A comparison among likelihood ratio, Lagrange Multiplier, and Wald tests. *Multivariate Behavioral Research*, *25*, 115-136.

Chou, C. –P., & Bentler, P. M. (1993). Invariant standardized estimated parameter change for model modification in covariance structure analysis. *Multivariate Behavioral Research, 28*, 97-110.

Cliff, N. (1983). Some cautions concerning the application of causal modeling methods. *Multivariate Behavioral Research*, *18*, 115-126.

Jorgensen, T. D. (2017). Applying permutation tests and multivariate modification indices to configurally invariant models that need respecification. *Frontiers in Psychology*. doi.org/10.3389/fpsyg.2017.01455

Kaplan, D. (1989). Model modification in covariance structure analysis: Application of the expected parameter change statistic. *Multivariate Behavioral Research*, *24*, 285-305.

Lee, S. –Y. (1985b). On testing functional constraints in structural equation models. *Biometrika*, *72*, 125-131.

Lee, S. –Y., & Tsang, S. Y. (1999). Constrained maximum likelihood estimation of two-level covariance structure model via EM type algorithms. Psychometrika, 64, 435-450.

Luijben, T. C., & Boomsma, A. (1988). Statistical guidance for model modification in covariance structure analysis*. Compstat*, 335–340.

MacCallum, R. C. (1986). Specification searches in covariance structure modeling. *Psychological Bulletin*, *100*, 107-120.

MacCallum, R. C., Roznowski, M. & Necowitz, L. B. (1992). Model modification in covariance structure analysis: The problem of capitalization on chance. *Psychological Bulletin*, *111*, 490-504.

Marcoulides, K. M., & Falk, C. F. (2018). Model specification searches in structural equation modeling with R. *Structural Equation Modeling*, *25*, 484-491.

Rao, C. R. (1948). Large sample tests of statistical hypotheses concerning several parameters with application to problems of estimation. *Proceedings of the Cambridge Philosophical Society*, *44*, 50-57.

Rindskopf, D. (1984b). Using phantom and imaginary latent variables to parameterize constraints in linear structural models. *Psychometrika*, *49*, 37- 47.

Saris, W. E., Satorra, A., & Sörbom, D. (1987). The detection and correction of specification errors in structural equation models. In C. Clogg (Ed.), *Sociological methodology 1987* (pp. 105-129). San Francisco: Jossey Bass.

Satorra, A. (1989). Alternative test criteria in covariance structure analysis: A unified approach. *Psychometrika*, *54*, 131-151.

Sörbom, D. (1975). Detection of correlated errors in longitudinal data. *British Journal of Mathematical and Statistical Psychology*, *28*, 138-151.

Sörbom, D. (1989). Model modification. *Psychometrika*, *54*, 371-384.

Tang, M. –L., & Bentler, P. M. (1998). Theory and method for constrained estimation in structural equation models with incomplete data. *Computational Statistics & Data Analysis*, *27*, 257-270.

Whittaker, T. A. (2012). Using the modification index and standardized expected parameter change for model modification. *Journal of Experimental Education*, *80*, 26-44.