RR – 85-03

HIERARCHICAL PATH MODELS WITH RANDOM COEFFICIENTS

Jan de Leeuw
Department of Data Theory FSW/RUL
Middelstegracht 4
2312 TW Leiden
The Netherlands

Abstract

In this short paper we extend the random coefficient regression model for multilevel analysis, that has been proposed recently by Mason, Wong & Entwisle and by De Leeuw & Kreft, to a random coefficient path model. We show that statistical and computational results derived for the earlier model can still be applied, provided the path coefficients of different endogeneous variables are independent.

Keywords

Multilevel analysis, random coefficient models, path analysis.

INTRODUCTION

A hierarchical regression model has two or more regression levels. The regression coefficients from a particular level are the dependent variables for the regression on the next higher level. Such hierarchical models have many applications in the social and behavioural sciences. A particulary important application in sociometry and in educational research is multilevel analysis, in which the levels of the hierarchy corresponds with different levels of aggregation of a set of units. Levels can be, for instance, pupils → classes → schools → neighbourhoods → towns → countries, etcetera. For a more thorough discussion we refer to Streitberg (1978).

Appropriate statistical models for multilevel analysis have not been considered until quite recently. Streitberg (1978) presents a linear model which takes the multilevel structure into account. He assumes that the dispersion matrix of the residuals has a particular interclass-correlation form. But this implies that his models are not hierarchical, because he does not really model the higher levels explicitly. Several authors have emphasized recently, that explicit modelling of both (or all) levels is possible by using random coefficient regression models (Tate and Wongbundhit, 1983, Mason, Wong, and Entwisle, 1983, De Leeuw and Kreft, 1985). Such random coefficient regression models have been studied fairly extensively in econometrics (Swamy, 1971, Rosenberg, 1973, Spjøtvoll, 1977, Mundlak, 1978, Johnson, 1977, 1980, Chow, 1984, Johansen, 1984). In statistics very similar models have often been presented in a Bayes or empirical Bayes context (Lindley and Smith, 1972, Smith, 1973, Polasek, 1984).

The general multilevel model of Mason, Wong, and Entwisle

(1983) and of De Leeuw and Kreft (1985) is a genuine two-step regression model. To present it we adopt the convenient formalism of underlining random variables (Hemelrijk, 1966). For individuals from group j we assume $\underline{y}_j = X_j \underline{b}_j + \underline{\varepsilon}_j$, with $\underline{\varepsilon}_j \sim N(0,\sigma_j^2 I)$. This defines the micro or within-group structure. The macro or between-group structure is defined by $\underline{b}_j = \Theta z_j + \underline{\delta}_j$, with $\underline{\delta}_j \sim N(0,\Omega)$. In the macro-equations Θ contains regression coefficients, while z_j are the measurements on the macro-level regressors for group j. If we combine the two levels into a single model we find $\underline{y}_j = X_j \Theta z_j + \underline{\zeta}_j$, where $\underline{\zeta}_j \sim N(0, X_j \Omega X_j' + \sigma_j^2 I)$. This is the 'reduced' form of the model which is used in the actual computations. Both Mason, Wong, and Entwisle (1983) and De Leeuw and Kreft (1985) discuss one-step and two-step ordinary least squares methods, the Swamy-Rao weighted least squares method, MINQUE estimates, and maximum likelihood estimates for this multilevel model.

Although the model discussed above is certainly interesting, it has two build-in restrictions of generality. In the first place there is only one micro-level dependent variable, and in the second place both micro-level and macro-level independent variables are considered as fixed. The first restriction is not a very serious one, because in principle extension to a multivariate linear model is straightforward. The second restriction, however, seems a rather more serious limitation. It has been argued by De Leeuw and Kreft (1985) that in educational research the assumption of fixed regressors at the micro-level is not very appropriate, because the sampling interpretation connected with fixed independent variables is often not realistic. The independent variables often have the same sampling characteristics as the dependent variables, and it seldom makes sense to assume that they are measured without error. Thus we need to get closer to the framework of structural covariance analysis, in which all variables are random and in which the distinction between dependent and independent variables is replaced by a more flexible path model.

In this paper we propose a general multilevel path model, and we indicate that the multilevel regression model of Mason, Wong and Entwisle (1983) and Leeuw and Kreft (1985) is a very special case. Nevertheless we also show that computational techniques used for the earlier multilevel model can be used almost without modification in this more general path model. We shall not go into the many possibilities offered by our more general approach, but we shall indicate some promising avenues for further research.

PATH MODEL

We present our results in the context of the following simple path model. Possible generalizations will be indicated below. Suppose $\underline{y}_{ij}$ $(i=1,\ldots,n;j=1,\ldots,m)$ are observable random variables, and $\underline{a}_{j\ell}$ $(j=1,\ldots,m;$ $\ell=1,\ldots,m;j>\ell)$ are unobservable or latent variables. We assume that for $j > 1$ we have a conditional density of the form

$$p(y_j|y_1,\ldots,y_{j-1},A) = N(Y_j a_j,\sigma_j^2 I). \tag{1}$$

Thus the conditional densities are spherical normal, with expectation $Y_j a_j$. Here Y_j is then n x (j - 1) matrix containing the 'previous' observed variables, and a_j is row j of A (more precisely the first j - 1 elements of that row, the others are zero). We also assume that

$$p(y_1|A) = N(0,\sigma_1^2 I), \tag{2}$$

and we assume that the rows of A are independent normals, with

$$p(a_j) = N(\alpha_j, \Omega_j). \tag{3}$$

We integrate out the a_j, and find the conditional densities

$$p(y_j | y_1, \ldots, y_{j-1}) = N(Y_j\alpha_j, Y_j\Omega_jY_j' + \sigma_j^2 I). \tag{4}$$

Equation (4) is our main result. It shows that factorization (1), which characterizes recursive path models, with independent errors, carries over to the random coefficient path model, provided the coefficients corresponding with different variables are independent.

The model given above is of some interest, perhaps also outside the multilevel context. We briefly indicate some possible generalizations. In the first place we can suppose, without further ado, that the exogeneous variable $\underline{y}_1$ is multidimensional. This merely changes (2) to $p(y_1 | A) = N(0, \Sigma)$. In the second place each of the y_j could be a block of variables. This complicates the notation, but not the theory, which becomes a straightforward generalization of theory for (fixed parameter) block recursive models. Thirdly there may be errors in variables, i.e. our theory has to be extended to LISREL or PLS models. This will require some additional effort, it is not merely a matter of notation. In the fourth place we must keep various special cases in mind. If all Ω_j are zero, for instance, the model becomes a fixed coefficient model. If we want certain additional regression coefficients to be zero (as in incomplete recursive path models) we set corresponding $\alpha_{j\ell}$ equal to zero, together with the corresponding diagonal element of Ω_j. Of course we can also set $\alpha_{j\ell} = 0$ without assuming in addition that $\omega_{j\ell\ell} = 0$. All these hypotheses are, at least in principle, testable by standard methods. A fifth generalization that must be mentioned assumes that in addition to the $\underline{a}_j$ also the $\underline{\sigma}_j^2$ are random variables. Such refinements can be incorporated using theory presented by Aragon (1985). And finally we must consider the possibility that the rows of $\underline{A}$ are correlated. This will make the theory much more complicated, but in some cases more realistic. Results on 'transmitted errors', discussed by Mundlak (1978), fit into this framework.

MULTILEVEL ASPECTS

The path model (4) is derived for a single population, and a sample from this population. If we have more than one population, we can use similar models for each population, and each population has its own parameters α_j, Ω_j, and σ_j^2. This does not really change anything, but matters become more interesting if the parameters in the different populations are related. In multilevel models, for instance, we use restrictions on the parameters to specify a model for the macro-level. Using the index $v = 1,\ldots,N$ for populations (i.e. schools, neighborhoods, etc), we impose the restrictions $\Omega_{jv} = \Omega_j$ and we want the α_{jv} to be a linear function of some fixed variables describing the macro-level units (populations). With these restrictions, and in the special case $m = 2$, path model (4) becomes identical to the multilevel model of Mason, Wong, and Entwisle (1983) and of De Leeuw and Kreft (1985). The $\underline{a}_{jv}$ ($v=1,\ldots,N$) are independent, and have distribution $N(\theta z_{jv},\Omega_j)$. It is clear that our more general approach offers many more possibilities for modelling on a macro-level, certainly if we take the possible generalizations from the previous section into account.

At the same time we see precise nature of multilevel models of this type somewhat more clearly. The issue of random coefficients and/or random regressors is quite irrelevant for the general idea. We have macro-levels, which define populations from which the micro-level units are sampled. The micro-models are defined for each population separately. They have parameters, and the macro-model is actually a model for these parameters. Of course the approach can be extended to the case in which there are more than two levels. This defines the so-called hierarchical models, which have been discussed by Lindley and Smith, Leamer, Polasek, and others. The models have both frequentist and Bayesian interpretations. But both for Bayesians and for frequentists they are most appropriate if the sampling framework consists of simple random samples from fixed populations. Other multilevel models will be necessary for cluster samples, and multi-stage samples.

ESTIMATION

We can be very brief about estimation methods. The complete range of procedures discussed by Mason, Wong, and Entwisle (1983) and by De Leeuw and Kreft (1985) applies without any modification, at least if the parameters of the m 'single equations' (4) are neatly separated. The only restrictions on the parameters must be 'within-equations', there must be no restrictions 'between-equations'. This is similar to the requirement that the rows of $\underline{A}$ must be independent. Of course 'between-population' restrictions are allowed, in fact they are even essential to make multilevel analysis possible. But given the separation of the equations we can do one-step and two-step unweighted least squares, we can compute the Rao-Swamy weighted least squares estimates, MINQUE estimates, and we can compute maximum likelihood estimates in various ways. For the details we refer to the publications mentioned earlier in this section, and to the references given in these publications.

REFERENCES

Aragon, Y. (1985), Random variance linear models: estimation, Computational Statistics Quarterly, in press.

Chow, G.C. (1984), Random and changing coefficient models, in Z. Griliches and M.D. Intriligator (eds), Handbook of Econometrics II, Amsterdam, North Holland Publishing Company.

De Leeuw, J. & Kreft, G.G., Random coefficient models for multilevel analysis, submitted for publication, 1985.

Hemelrijk, J. (1966), Underlining random variables, Statistica Neerlandica, 20, 1-8.

Johansen, S. (1984), Functional relations, random coefficients, and nonlinear regression with application to kinetic data, New York, Springer-Verlag.

Johnson, L.W. (1977), Stochastic parameter regression: an annotated bibliography, International statistical Review, 45, 257-272.

Johnson, L.W. (1980), Stochastic paramter regression: an additional annotated bibliography, International Statistical Review, 48, 95-102.

Lindley, D.V. & Smith, A.F.M. (1972), Bayes estimates for the linear model (with discussion), Journal of the Royal Statistical Society, B34, 1-41.

Mason, W,M., Wong, G.Y., & Entwisle, B. (1983), Contextual analysis through the multilevel linear model, in S. Leinhardt (ed), Sociological methodology 1983/1984 San Francisco, Jossey-Bass.

Mundlak, Y. (1978), Models with variable coefficients: integration and extension, Annales de l'INSEE, 30-31, 483-509.

Polasek, W. (1984), Multivariate regression systems, in J.B. Kadane (ed), Robustness of Bayesian analysis, Amsterdam, North Holland.

Rosenberg, B. (1973), A survey of stochastic parameter regression, Annals of economic and social measurement 2, 381-397.

Smith, A.F.M. (1973), A general Bayesian linear model, Journal of the Royal Statistical Society, B35, 67-75.

Spjøtvoll, E. (1977), Random coefficient regression models: a review. Mathematische Operationsforschung und Statistik, series Statistics, 8, 69-93.

Streitberg, B. (1978), Schätzung von Kovarianzstrukturen in linearen Zwei- und Mehrebenenmodellen, Unpublished Doctoral Dissertatiun, Free University of Berlin.

Swamy, P.A.V.B. (1971), Statistical inference in random coefficient regression models, New York, Springer-Verlag.

Tate, R.L. & Wongbundhit, Y. (1983), Random versus nonrandom coefficient models for multilevel analysis, Journal of Educational Statistics, 8, 103-120.