CHAPTER 19

Foundational Issues in Intraindividual Longitudinal Analysis

MICHAEL J. ROVINE
LAWRENCE L. LO

Central Issues

Single-subject and person-specific approaches to the modeling of developmental processes are becoming more central in the description of complex and time-varying processes. These methods, which have a relatively long tradition in psychology (Kratochwill, 1978), have been labeled **idiographic** (Lamiell, 1981; Zevon & Tellegen, 1982) and have been used to describe intraindividual patterns of change (Nesselroade & Ford, 1985). The more general behavioral laws described through cross-sectional studies of a sample of individuals have been labeled **nomothetic**. Most researchers are aware of the problems inherent in interpreting results from cross-sectional data (Baltes, Reese, & Nesselroade, 1977). Many, however, are much less aware of the interpretive problems that arise when the data collected represent a few repeated occasions on a sample of individuals. Molenaar (2004) has described many of these problems and has indicated how these difficulties can often be resolved only with time series data. Although such data have long been collected in a number of different research areas (Jones & Nesselroade, 1990; Luborsky & Mintz, 1972), these data are often not available because of the difficulty and expense involved in collecting a large number of occasions. Fortunately, due to new technologies, time series data are becoming much easier and less expensive to collect, and methods for estimating many different models are being used more often (Walls & Schafer, 2006).

The availability of time series data (intensive data collected on each individual) through diaries, personal data assistants, and "smartphone" devices through measures of
physiological characteristics (EKG, fMRI, heart rate, blood pressure, etc.), and by other means have brought time series approaches into the domain of the developmental sciences (Molenar, 1984). Time series approaches, including Box–Jenkins type autoregressive integrated moving average (ARIMA) models (Box & Jenkins, 1970), multivariate vector autoregressive models (Lutkepohl, 2005), P-technique factor models and dynamic factor models (Molenar, 1985; Wood & Brown, 1994), and state-space models (Molenar, 1994), among others, have been appearing more often in the developmental literature. Both time domain (Box & Jenkins, 1970) and frequency domain (Jenkins & Watts, 1968) approaches developed for engineering, econometric, and medical applications are now becoming important tools in the social sciences. As with the switch to any new set of methods, a certain amount of reluctance and misunderstanding exists. In this chapter we wish to discuss the following issues involved in these important approaches.

**Person-Specific versus Group-Based Models**

With separate models estimated for a set of individuals, it becomes possible to test whether a single model adequately describes the group. When only a single model is estimated based on the entire group’s data, this becomes impossible. The assumption of such group-based models is that there is a common model that holds equally well for each member of the sample. By comparing separately estimated individual models with the group model using time series data, we can specifically test whether the individuals can be adequately described by a group model.

Because there is very often a fine line between a group-oriented approach and a person-specific approach, we review a number of different currently popular methods and describe the degree to which they are person specific. Because many models that generate measures of individual differences are based on parameters estimated on the group, we distinguish between these models (group based) and models in which parameters are estimated separately for each individual (person specific).

**Time Series Analogues of Group-Based Models**

Many different models for individual differences exist, some of which are person specific based on time series data and some of which provide individual descriptive statistics that are nevertheless based on group models. Two important modeling approaches that fall into the latter category are the lagged regression and path analytic approaches and the multilevel growth curve model and repeated-measures analysis of variance (ANOVA) approaches. Although these models generate descriptive statistics that can be used to describe individual differences, they are really group models that allow post hoc predictions of these individual differences. As a result, they provide limited person-specific information.

**Single-Subject Models Based on Cattell’s P-Technique Notion**

Because the tradition of person-specific modeling comes, in part, from the assessment of interindividual differences in intraindividual change, we describe how the idea of the data box (Cattell, 1966) led to a family of approaches that come under the heading of P-technique analysis. By adding lagged variables to the P-technique data, we add a number of additional

analytical possibilities, including longitudinal factor analysis, vector-autoregressive models, dynamic factor models, and state-space models.

**What Constitutes a Sample**

An important difference between group-based and person-specific analyses is the idea of a sample. In a group-based analysis, the parameters of the model are often used to generalize to a population of interest. If the sample is representative of the population, one can say something important about the population. In a person-specific analysis, the generalization is related to the individual model. Because time is essentially taking the place of number of participants, the generalization is made across the time dimension. Most typically, different models will arise for different individuals. Subgroups of individuals can be constructed by pooling those individuals whose models are identical. Nelson and Molenar (1999) described methods for pooling (as opposed to sampling). As a simple example, P-technique factor analysis requires a multivariate time series collected on an individual. For more than one individual, the multiple P-technique analyses are akin to multiple group structural equation models (SEM). By constraining solutions equal across individuals we can test whether the same model holds for all.

**Recursive Estimation and the State-Space Model**

Estimation methods for many of these models fall into two general categories: complete data procedures that estimate parameters once all of the data have been collected and recursive estimation procedures in which parameters of the model are updated sequentially as each new observation appears. Once all of the data have arrived, the recursive and complete data estimates of the same statistical model typically result in the same parameter estimates. However, for certain types of problems, recursive estimation simplifies the computations. For problems in which the model changes as the data arrive, recursive procedures allow certain time-varying characteristics, including the possibility of estimating time-varying parameters in ways that would be difficult using the complete data approach. A very general model for describing recursive estimations is the state-space model (Harvey, 1989). We show a basic form of this model and suggest how it can be used to approach the solution to certain estimation problems.

Recursive estimation arguably began with Gauss and then disappeared until being rediscovered by Plackett (1950). It was generalized by Kalman as the famous Kalman filter (Kalman, 1960; Kalman & Bucy, 1961), which represented a general solution for estimating the state in the state-space model formulation. The filtering approach not only allowed the opportunity for estimation of time series parameters but also introduced the possibility of control; namely, having targets for certain variables in the model and changing other variables in the model to move toward the optimal target. Although the derivation of the filter by Kalman is based on a somewhat difficult area of mathematics, the filter can be described using regression concepts. To give a flavor of the recursive method of estimation, we show how regression models can be reconstituted in recursive form. The resulting regression equations have the form of a filter. Using the pieces that compose these equations, we can indicate the relationship between the recursive regression estimator and the Kalman filter.
P-Technique and Dynamic Factor Models

The Data Box

Cattell (1966) presented a schematic for describing data from a multivariate study. As shown in Figure 19.1, the data box has variables, individuals, and occasions, creating a three-dimensional structure. Each two-dimensional slice of the box represents a way of constructing an association matrix. The vertical axis indicates the dimension across persons. The horizontal planes contain axes that identify the dimension across variables (upper left to lower right) and the dimension across occasions (lower left to upper right). Each two-dimensional slice provides an association matrix for one dimension that pools across the other dimensional component. Cattell gave each two-dimensional facet a letter name. For example, the individuals-by-variables facet (pooled across individuals) was called R, and the factor analysis resulting from analyzing the variable-by-variable association matrix was called R-technique; this is the most common form of factor analysis. Although all facets were given names, some were more successful in generating interpretable research results than others. Important for studies of intraindividual differences was the P facet, which created an occasion-by-variable rectangular data set for a single subject (pooled across occasions). Factor analysis of this type of data set was referred to as P-technique. The R-technique (interindividual-vertical) and P-technique (intraindividual-horizontal) slices are represented in Figure 19.1.

P-Technique Factor Analysis

Research in which P-technique factor analysis has been used includes studies of stepchildren's emotional experiences (Corneal & Nesselroade, 1994), mood change (Bath, Daly, &

\[ y_t = \Lambda \eta_t + \varepsilon_t \]  

where \( y_t \) is a vector of observed variables at time \( t \), \( \Lambda \) is a matrix of factor loadings, \( \eta_t \) is a vector of latent variables at time \( t \), and \( \varepsilon_t \) is a vector of residuals. The data set for the P-technique analysis is the horizontal rectangular data set highlighted in Figure 19.1. Pooling across occasions results in a variable-by-variable association matrix.

With multiple participants, we can estimate separate P-technique solutions. We can then compare the solutions to see whether the same model tends to describe different individuals. If that is the case, the separate data sets could be pooled to calculate a single group analysis. Most typically, that is not the case. As an example, in Table 19.1 we show two different P-technique factor analyses from a study of mood in which the Positive and Negative Affect Schedule (PANAS) was administered to stepchildren after an interaction with their stepparents (Rovine, Molenaar, & Cornel, 1999). The factor patterns show interesting and idiosyncratic patterns. Notice that factors include combinations of positive and negative items and that the factor correlations suggest that these factors are relatively independent. This contradicts the typical group results for the PANAS, which tends to produce positive and negative subscales. The interesting individual differences tend to be averaged out of such pooled analyses. In the event that certain subsets of the sample have structurally identical models, methods are available for pooling the results of the individual P-technique analyses (Nesselroade & Molenaar, 1999).

Although P-technique was relatively accepted as an important method for describing multivariate time series, a number of criticisms emerged (Anderson, 1963; Holtzman, 1982). Most notably, no lagged values of the variables appear in the data matrix. A lagged value or lagged relationship is the relationship between variables at two time points, where the degree of lag indicates the gap in time; for example, a lag 2 is the relationship between a variable at time \( t \) and the variable at time \( t - 2 \). For example, if we are measuring mood and satisfaction on a daily basis and today's mood predicts tomorrow's satisfaction, that is a lag 1 relationship. If today's mood predicts satisfaction 2 days later, that is a lag 2 relationship. If today's mood predicts today's satisfaction, that is a simultaneous relationship. P-technique models only simultaneous relationships among the data. Without the lags the timing of the components may be lost. Effects resulting from this omission were first presented by Anderson (1963). More recently Molenaar and Nesselroade (2009) demonstrated the relative robustness of the P-technique solution even in the presence of lagged relationships.

Going beyond P-technique, Molenaar (1985) presented a very general model, the dynamic factor model, that properly addressed the question of lags and the order of the lagged relationships. We can add lags to the P-technique data set by adding an offset version of each variable as an additional variable. In Figure 19.2 we add the first-order lags for each variable in the data set.
### IV. INTRAINDIVIDUAL LONGITUDINAL ANALYSIS

#### TABLE 19.1. P-Technique Factor Analyses for Stepsons A and B

<table>
<thead>
<tr>
<th>Factor loadings</th>
<th>Stepson A</th>
<th>Stepson B</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Interest</td>
<td>.64</td>
<td></td>
<td>.77</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Excited</td>
<td>.67</td>
<td></td>
<td>.81</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Enthused</td>
<td>.72</td>
<td>-.38</td>
<td>.81</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inspired</td>
<td>.42</td>
<td>-.36</td>
<td>.56</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alert</td>
<td>.66</td>
<td>.34</td>
<td>.77</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Attentive</td>
<td>.74</td>
<td>.33</td>
<td>.58</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Satisfied</td>
<td>.62</td>
<td>-.59</td>
<td>.58</td>
<td>.61</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Proud</td>
<td>.66</td>
<td>-.39</td>
<td>-.60</td>
<td>.65</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Distressed</td>
<td>.85</td>
<td></td>
<td>.96</td>
<td>.92</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Upset</td>
<td>.87</td>
<td>-.23</td>
<td>.92</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Strong</td>
<td>.51</td>
<td>.30</td>
<td>.23</td>
<td>.56</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hostile</td>
<td>.91</td>
<td>.29</td>
<td>.79</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Irritable</td>
<td>.88</td>
<td>-.32</td>
<td>.92</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Discouraged</td>
<td>.74</td>
<td>.33</td>
<td>.24</td>
<td>.82</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Accepting</td>
<td>.46</td>
<td>-.63</td>
<td>.65</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NotWant</td>
<td>.97</td>
<td></td>
<td>.74</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Deterred</td>
<td>.51</td>
<td>.34</td>
<td>.25</td>
<td>.75</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contented</td>
<td>.32</td>
<td>-.46</td>
<td>-.55</td>
<td>.57</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Scared</td>
<td>.71</td>
<td></td>
<td>.79</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nervous</td>
<td>.92</td>
<td></td>
<td>.78</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jittery</td>
<td>.45</td>
<td>.73</td>
<td>.27</td>
<td>.81</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Afraid</td>
<td>.98</td>
<td>-.13</td>
<td>.79</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Humiliated</td>
<td>.35</td>
<td>-.49</td>
<td>.57</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Loved</td>
<td>.43</td>
<td></td>
<td>.69</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Active</td>
<td>.43</td>
<td>.37</td>
<td></td>
<td></td>
<td>.81</td>
<td></td>
</tr>
<tr>
<td>Liking</td>
<td>.62</td>
<td></td>
<td>.72</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ashamed</td>
<td>.49</td>
<td>-.41</td>
<td>.82</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Close</td>
<td>.73</td>
<td></td>
<td>.68</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Factor intercorrelations</th>
<th>Interest</th>
<th>Anger</th>
<th>Anxiety</th>
<th>Affection</th>
<th>Closeness</th>
<th>Interest</th>
<th>Anger</th>
<th>Strength</th>
</tr>
</thead>
<tbody>
<tr>
<td>Interest</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anger</td>
<td>.36</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td>-.10</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>Anxiety</td>
<td>-.06</td>
<td>-.08</td>
<td>1.00</td>
<td></td>
<td></td>
<td>.42</td>
<td>.65</td>
<td>1.00</td>
</tr>
<tr>
<td>Affection</td>
<td>.13</td>
<td>-.03</td>
<td>.11</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Closeness</td>
<td>.37</td>
<td>-.36</td>
<td>.14</td>
<td>.23</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

#### Dynamic Factor Analysis

Molenaar's (1985) dynamic factor model represents a generalization of the P-technique model. Although particular versions of this model had been previously discussed by Priestley, Subba Rao, and Tong (1973), Brillinger (1975), and others, Molenaar provided a very general method that, among other things, translated nicely to the domain of structural equation modeling.

The dynamic factor model is

\[ y_t = \sum_{\alpha=0}^{s} \Lambda_{\alpha} \eta_{t-\alpha} + \epsilon_t \]  

(19.2)

where \( \Lambda_{\alpha} \) contains the factor loading for lag, \( \alpha \). With this model the researcher can determine the degree to which latent variables and residuals are correlated over time. The P-technique model is a special case of the dynamic factor model with \( s = 0 \) (i.e., only a contemporaneous factor loading matrix).

Because many find the model relatively complex, we think that it is helpful to consider a path diagram of a simple dynamic factor model. This helps to show how the addition of lags adds a set of required constraints to the model. In Figure 19.2 we show a dynamic factor for five observed variables with one factor and a lag order of 1 based on Wood and Brown (1994). The model includes variables at times \( t \) and \( t-1 \). The model represents the fact that at any particular occasion, the observed variables are dependent on both a "today" factor and a "yesterday" factor (Wood & Brown, 1994). To also allow the \( t-1 \) observed variables to be accounted for by a previous day, we also add a \( t-2 \) latent variable. Notice that the observed variables are predicted by the concurrent and previous latent variables. No regression relationship is included between the lagged latent variables. Correlations among the residuals are included for the current and previous occasion variable pairs.

Although the model looks very similar to the more standard factor model, its inclusion of lags adds some additional technical complications to the estimation of the model. An ordinary covariance matrix is not the proper input matrix. The model requires a special type of input matrix (the block Toeplitz matrix) of simultaneous and lagged covariances. A routine for properly estimating the matrix can be found in Wood and Brown (1994).

The model requires the set of constraints indicated in Figure 19.2. Because of the redundancy of the model, fit indices for the model must be calculated by hand. These and other technical problems involved in estimating the model are discussed in Wood and Brown (1994).

![Figure 19.2. Dynamic factor model with lagged and concurrent factor loadings.](image)
Molenar, Sinclair, Rovine, Ram, and Conneal (2009) discuss situations in which the P-technique assumptions do not hold. Although the general dynamic factor model (Molenar, 1985) would represent a proper alternative with sufficient data, that model requires a long time series to identify the proper order of the lagged relationship and to allow the computation of a stable model given the number of variables.

**Person-Specific Individual Difference versus Group Models**

A very subtle difference can exist between group-based models that generate statistics that describe differences in individuals and genuine person-specific models. The simplest way to make this distinction is to consider whether the parameters of a model are estimated based on the group or whether a separate model for each individual is used to generate the parameters. At the heart of the matter is what constitutes a subgroup. For the group-based models, a subgroup represents an a priori definition based on some known variable or, in the case of latent class or mixture models, a certain similarity based on group distributional characteristics. For the person-specific model, a subgroup represents a set of individuals who can be shown empirically to have identical or very similar model parameter values. Although both modeling approaches can provide useful information, the group-based model comes with one important assumption: namely, that the same model holds for all members of the sample.

The *ergodicity theorem* (Birkhoff, 1931) implies that for a result from a group-based model to properly apply to an individual, the group model has to be the appropriate model for each individual. In the absence of time series data, this is an untested assumption. Molenar (2004) has shown dramatically how factor models derived from cross-sectional analysis on a group covariance matrix can have little to do with the models that best describe each individual. Using a data set provided by Borkenhagen and Ostendorf (1998), Molenar has demonstrated the implications of violations of ergodicity for personality factors based on the Big Five, namely, for individuals, the Big Five seems to be a poor model. It is only when pooling across individuals that the five factors emerge. Rovine, Molenar, and Conneal (1999; Molenar, Rovine, & Conneal, 1999) showed a similar result based on time series assessments of mood, showing that the positive and negative factors of the PANAS may be an artifact of pooling.

The possible violation of the ergodicity assumption always lurks in the background in any study using either cross-sectional or short longitudinal data. Without either prior time series results validating the group model for individuals or concurrently collected time series data, this concern must at least be considered.

Other modeling approaches attempt to account for heterogeneity of intraindividual variation. Latent class growth analysis (Nagin, 1999) and growth mixture modeling (Muthén & Shedden, 1999) address group heterogeneity by extraction of subgroups that differ in structure. Although these methods allow heterogeneity, within an aggregate group they still assume homogeneity within subgroups (Sterba & Bauer, 2010). Consequently, these modeling approaches both make assumptions of ergodicity.

To show how the individual difference measures are generated from a group-based model, we will consider two commonly used models: the linear growth curve model (McArdle & Epstein, 1987) and the lagged regression or panel analysis model (Blalock, 2007; Little, Preacher, Selig, & Card, 2007).

### A Linear Growth Curve Model

There are a number of equivalent ways to define the linear growth curve model (Bryk & Raudenbush, 1992; Goldstein, 1995; McArdle & Epstein, 1987; Rovine & Molenar, 2005). One useful way is based on the linear mixed model (Laird & Ware, 1982). The linear mixed model can be written as

\[
\mathbf{y} = \mathbf{X} \boldsymbol{\beta} + \mathbf{Z} \mathbf{\gamma} + \mathbf{\epsilon},
\]

(19.3)

where \(\mathbf{y}\) is a vector of repeatedly measured dependent variable values; \(\mathbf{X}\) is the design matrix (matrix of predictor variable values related to the fixed effects); \(\mathbf{\beta}\) is the vector of fixed effect parameters; \(\mathbf{Z}\) is the design matrix related to the random effects; \(\mathbf{\gamma}\) is the vector of random effects parameters; and \(\mathbf{\epsilon}\) is the vector of residuals. \(\mathbf{Z}\), \(\mathbf{\gamma}\), and \(\mathbf{\epsilon}\) can vary among individuals.

The linear growth curve model estimates a set of fixed effects parameters that define straight lines for the different groups. By including polynomial terms (powers, centered powers, or orthogonal polynomials), curves of any shape can be modeled. For two groups with four repeats, the fixed effects part of the model in matrix form becomes

\[
\begin{bmatrix}
\mathbf{y}_{11} \\
\mathbf{y}_{12} \\
\mathbf{y}_{13} \\
\mathbf{y}_{14}
\end{bmatrix} = 
\begin{bmatrix}
1 & 0 & 0 & 0 \\
1 & 1 & 0 & 1 \\
1 & 2 & 0 & 0 \\
1 & 3 & 0 & 0
\end{bmatrix}
\begin{bmatrix}
\beta_1 \\
\beta_2 \\
\beta_3 \\
\beta_4
\end{bmatrix} + 
\begin{bmatrix}
\gamma_{11} \\
\gamma_{12} \\
\gamma_{13} \\
\gamma_{14}
\end{bmatrix} + 
\begin{bmatrix}
\epsilon_{11} \\
\epsilon_{12} \\
\epsilon_{13} \\
\epsilon_{14}
\end{bmatrix}
\]

Group 1

\[
\begin{bmatrix}
\mathbf{y}_{21} \\
\mathbf{y}_{22} \\
\mathbf{y}_{23} \\
\mathbf{y}_{24}
\end{bmatrix} = 
\begin{bmatrix}
1 & 0 & 1 & 0 \\
1 & 1 & 1 & 1 \\
1 & 2 & 1 & 2 \\
1 & 3 & 1 & 3
\end{bmatrix}
\begin{bmatrix}
\beta_1 \\
\beta_2 \\
\beta_3 \\
\beta_4
\end{bmatrix} + 
\begin{bmatrix}
\gamma_{21} \\
\gamma_{22} \\
\gamma_{23} \\
\gamma_{24}
\end{bmatrix} + 
\begin{bmatrix}
\epsilon_{21} \\
\epsilon_{22} \\
\epsilon_{23} \\
\epsilon_{24}
\end{bmatrix}
\]

Group 2

(19.4)

where \(\beta_1\) and \(\beta_2\) are the respective intercept and slope for the first group and \(\beta_3\) and \(\beta_4\) are the respective differences in intercept and slope between the two groups. These parameters are fixed for the sample. \(\gamma_1\) and \(\gamma_2\) are the random effects (in this case random slopes and intercepts). These are predicted post hoc based on the group parameters. Given by Henderson (1975), the fixed and random effects are

\[
\begin{bmatrix}
\beta_1 \\
\beta_2 \\
\beta_3 \\
\beta_4
\end{bmatrix} = 
\left(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X}\right)^{-1} \left(\mathbf{X}'\mathbf{V}^{-1}\mathbf{Y}\right)
\]

(19.5)

where \(\mathbf{V} = \text{E}(\mathbf{Z}\mathbf{\gamma} + \mathbf{\epsilon}X\mathbf{\gamma} + \mathbf{\epsilon},^\prime) = \mathbf{GZ}\mathbf{G} + \mathbf{R}\). \(\mathbf{G}\) is the covariance matrix of the random effects under the assumption
Thus the random effects are group-based estimates. This is quite opposed to the metaphorical notion that each person’s intercept and slope are estimated. They are post hoc predictions of the residuals based on the group model, not person-specific individual estimates.

The random effects typically described by investigators are measures of individual differences. These effects are basically regression residuals estimated under a specific set of assumptions for a particular regression model. In a growth curve setting, this becomes the requirement that each individual follow the same trajectory shape. Whereas the parameter values can differ from person to person, the parameterization of the curve typically does not. In situations in which the type of trajectory varies among individuals, the growth curve approach can misrepresent how individuals differ. In addition, the inference regarding the fixed effect parameters can be less than optimal when a single error model does not describe all individuals equally well (Liu, Rovine, & Molenaar, 2010).

The person-specific analogue to the growth curve model could require fitting a separate polynomial curve of an order to be empirically determined for each individual. With most typically designed studies, the number of occasions would be far too small to estimate parameters of this model with any precision. As a result, a growth curve along with its assumptions represents a necessary practical compromise.

**Lagged Regression Models**

Like the linear growth curve, the stability and lagged regression models are group-based models. The individual difference measures are actually post hoc predictions based on the group parameter estimates. Although this is perfectly acceptable for any question involving group differences, the question will always remain of whether the group model adequately describes any individual. Again, *ergodicity* remains an untested assumption.

**Moving to a True Individual Difference Regression Model**

To see how we can move to a true individual differences model, we can begin with a stability model for a variable, $x$, as described earlier. In Figure 19.3 we represent such a model for a single variable measured on five repeated occasions.

![FIGURE 19.3. Stability model for a single variable measured on five occasions.](image)

This figure represents a first-order autoregressive model (i.e., each occasion is regressed on only the prior occasion). To estimate the parameters of this model, we need a sample of individuals, $n$. The model is equivalent to estimating four separate regression equations with no restrictions on the parameters. We certainly could not estimate the model based on a single individual. However, with some modifications to that model, we can obtain parameter estimates for this model for a single subject. This alternative model is pictured in Figure 19.4. Here we extend the model from five occasions to $n$ occasions. We constrain the parameters between each occasion to be equal. We also assume the variance of the residuals at each occasion (the innovations) to be equal. (The innovation is the difference between the observed value at time $t$ and the predicted value based on prior information up to time $t$.) Because the regression coefficient for each pair of occasions is the same value, $\beta_{t,t-1}$, we can stack the data by lagging the variable, $x$, as described above. Regressing $y_t$ onto $y_{t-1}$ gives us the regression coefficient, $\beta_{t,t-1}$. This works because the regression of each occasion onto the previous occasion is the same value; so, wherever in the variable sequence one is, the prediction of the next occasion is always the same; hence the lagged variable. For a second-order autoregressive model, we would simply add $y_{t-2}$ as a second predictor. Given this strategy we can easily extend this logic to any lagged or cross-lagged regression model. Consider the regression of $y$ onto $x$ for a single subject shown in Figure 19.5. Again, the assumption that the regression coefficient is identical across occasions requires the lagged variable as a predictor and allows the model to be estimated. This model can be expressed as the set of autoregressive equations:

$$
\begin{align*}
y_{t+1} &= \beta_{11}y_t + \beta_{12}x_t + \epsilon_{t+1} \\
x_{t+1} &= \beta_{21}y_t + \epsilon_{2,t+1}
\end{align*}
$$

where $\beta_{11}$ is the stability coefficient for $y$ and $\beta_{12}$ is the lagged regression coefficient of $y$ onto $x$.

Using the same logic we can define a cross-lagged regression model (see Figure 19.6) that is the time series analogue to the standard panel cross-lagged regression model. This model can be expressed as the set of autoregressive equations:

$$
\begin{align*}
y_{t+1} &= \beta_{11}y_t + \beta_{12}x_t + \epsilon_{t+1} \\
x_{t+1} &= \beta_{21}y_t + \beta_{22}x_t + \epsilon_{2,t+1}
\end{align*}
$$

![FIGURE 19.4. First-order autoregressive model for a single variable measured on $n$ occasions.](image)

![FIGURE 19.5. Lagged regression model for time series data.](image)
Moving to a True Individual Difference Curve Model from a Trend Model

As we mentioned before, the growth curve model (either latent or manifest) is a group-based model in which individual difference measures are predicted post hoc based on the common model and a set of relatively restrictive assumptions. Typically, all individuals are assumed to follow the same general shape with individual differences represented as best linear unbiased predictions (BLUPs) of the individual curves. With time series data, unique trends or curve shapes can be estimated separately for individuals.

In time series modeling, we know that in order to properly model the covariance relationships among the variables in the series, we require at least a weakly stationary series. The variance function must be a constant, the series must be detrended, and the covariance function among the different series must also be constant. A thorough discussion of ergodicity and the problems of making inferences about individual processes from group-level analysis is presented by Molenaar (2004). Regarding detrending, there are basically two approaches. If we are not interested in modeling the trend in the series, we can difference the series, where differencing simply means taking the difference score between each adjacent pair of measures and treating these differences as a new series. Differencing once removes linear trends, differencing twice removes quadratic trends, and so forth. If we are interested in modeling the trends, we can fit a trend model to each time series. A simple approach would involve fitting polynomials (linear, quadratic, cubic, etc.) to each series until the residuals of the resultant series have a constant mean function. We can then determine the order of the process for the residuals. With an overall linear trend and a first-order autoregressive process, the equation for this model would be

$$y_{t+1} = \beta_0 + \beta_1 t + \beta_2 y_t + \epsilon_{t+1}$$ \hspace{1cm} (19.10)

As in all person-specific approaches we can peruse individual models and determine which individuals can be pooled together.

Recursive Estimation and the State-Space Model

Typically in statistical analysis with cross-sectional or longitudinal data with relatively few occasions, we estimate the model using all of the data. Often the study is over, all of the data are collected, and the time for analysis is nigh. In a time series study, however, some subset of the data might be available, but a fair amount of time may pass before all of the data are in. There might be circumstances in which it is important to analyze whatever data are available (e.g., how effective a medical treatment is) and to update the results as each new data point arrives. Such problems suggest recursive estimation over a complete data analysis.

Recursive estimation simply means reestimating a statistic as each new data point arrives. For data that are received over time, a recursive estimator allows the estimator of the parameter to include all of the data that are currently available and to use them to provide a relatively simple formula to update the estimate as each new data point is encountered. Once all of the data are in, the parameter estimate derived recursively will be the same as that estimate using complete data. However, for many types of problems, including problems that unfold in real time, the recursive approach and, in particular, the Kalman
filter has many advantages (Chow, Hamagami, & Nesselroade, 2007; Molenar et al., 2009). We present a short introduction to recursive estimation and show how certain estimation equations can be rewritten recursively. Using this approach to regression, we provide a hopefully intuitive link to the Kalman filter.

**Estimating a Mean Recursively**

To estimate a mean recursively we consider the data set in Table 19.2. Given the formula for the mean

$$\bar{y}_N = \sum_{i=1}^{N} y_i \left( \frac{1}{N} \right)$$

we can rewrite this as

$$\bar{y}_N = \sum_{i=1}^{N-1} y_i \left( \frac{1}{N} \right) + y_N \left( \frac{1}{N} \right) = \sum_{i=1}^{N-1} y_i \left( \frac{1}{N-1} \right) \left( \frac{N-1}{N} \right) + y_N \left( \frac{1}{N} \right)$$

$$= \bar{y}_{N-1} \left( \frac{N-1}{N} \right) + \frac{y_N}{N}$$

which gives the recursion relationship. Using the data in Table 19.2, we see that for line 6

$$\bar{y}_6 = \left( \frac{5}{6} \right) + \frac{9}{6} = \frac{24}{6} = 4$$

which is the mean up to observation 6. As expected, once all of the data are in, the mean is identical to the mean of all observations. If we think of the sequence of numbers as a time series, with each new observation representing a new occasion, we could plot the means by occasion and observe how the mean develops over time.

To begin the recursion, we need to know the value of the mean at \( N = 1 \), which is trivially the value of the first observation. Note, however, that we could arbitrarily pick a

| Table 19.2. Recursive Estimation of a Mean |
|-----------------|-----|-----|
| \( N \) | \( y \) | \( \bar{y} \) |
| 1 | 1 | 1 |
| 2 | 3 | 2 |
| 3 | 5 | 3 |
| 4 | 3 | 3 |
| 5 | 3 | 3 |
| 6 | 9 | 4 |
| 7 | 4 | 4 |
| 8 | 4 | 4 |

number. If the sample were small, the effect of picking this number could be dramatic on the overall mean. As the sample gets larger, the effect of this initial selection would diminish.

**Estimating a Simple Regression Recursively**

The recursion for the mean is relatively simple. For a linear regression, the problem becomes slightly more complex. To show one way of developing the recursion, we begin with a simple linear regression model. To additionally simplify the model, we assume that the data have been centered, so the intercept of the model is 0.

The equation for the simple regression model is

$$y_i = X_i \beta_1 + \epsilon_i \quad i = 1, \ldots, T$$

Given this equation, the regression weight is indexed by \( i \) to suggest that we are interested in estimating the regression for the first \( i \) observations.

Just using the first \( r \) observations, the regression coefficient is

$$\beta_r = \left( \sum_{i=1}^{r} x_i^2 \right)^{-1} \left( \sum_{i=1}^{r} x_i y_i \right)$$

Using algebraic manipulations, this equation can be rewritten as

$$\beta_r = \beta_{r-1} + \left( \sum_{i=1}^{r} x_i \right)^{-1} x_r (y_r - \beta_{r-1} x_r)$$

The multiple regression model is a straightforward extension of the simple regression model (Brown, Durbin, & Evans, 1975; Plackett, 1950). As is very often the case, we can write the multivariate version of an equation by taking the univariate version and replacing the scalars in the model with the appropriate vectors and matrices. Taking equation 19.14, the recursive estimation formula for the simple regression, we can extend it to the multiple regression recursion

$$b_r = b_{r-1} + (X_r^T X_r)^{-1} (X_r^T y_r - X_r^T b_{r-1})$$

The recursion typically includes a second recursion for the inverse of the sum of squares and cross-products (SSCP) matrix of the predictors, \((X_r^T X_r)^{-1}\). This recursion was originally presented by Plackett (1950).

**The State-Space Model**

The state-space model is a general recursion form that consists of two equations: the observation equation and the transition equation. Many variants of the basic model exist (Aoki, 1967; Durbin & Koopman, 2001; Harvey, 1989; Kim & Nelson, 1998; Simon, 2006), including continuous and discrete versions of the model. A good introduction to the concepts underlying the model and Kalman filter can be found in Maybeck (1979). Here we present
a simple form of the discrete-linear Gaussian state-space model. This can be expressed as two basic equations: the observation equation and the state or transition equation. A basic form of these equations (including no exogenous variables) includes the observation and state or transition equation.

\[ y_t = H x_t + u_t \quad \text{is the observation equation} \quad (19.16) \]

Following conventions we name the state vector \( x(t) \). Here \( y_t \) is a vector of observations. This equation has the form of a regression equation in which the parameter vector \( x(t) \) can vary over time and \( u(t) \) is a set of measurement errors. \( H \) is a design matrix that we will initially assume to be known.

\[ x(t + 1) = \Phi x(t) + w(t + 1) \quad \text{is the transition equation} \quad (19.17) \]

which is a first-order vector autoregressive model in which \( \Phi \) represents a set of regression weights linking elements of the state at time \( t \) to the same elements at time \( t + 1 \). \( w(t + 1) \) is the vector of innovations (differences between the predicted state and the actual state at time \( t + 1 \)). The error term in the observation equation corresponds to the "superimposed" error, while the innovation term in the transition equation corresponds to the disturbance error that was characterized by Yule. In the most general model, all parameter matrices, including \( H \) and \( A \), can be time varying.

To show a simple example of a state-space model, we can rewrite the second-order autoregressive model

\[ y_{t+1} = \beta_{1,t} y_t + \beta_{2,t} y_{t-1} + \epsilon_{t+1} \quad (19.18) \]

as

\[ y_t = \begin{bmatrix} 1 & 1 & 1 \end{bmatrix} x(t) \]

\[ \begin{bmatrix} y_{t+1} \\ y_t \\ y_{t-1} \end{bmatrix} = \begin{bmatrix} \beta_{1,t} & \beta_{2,t} & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} y_t \\ y_{t-1} \\ y_{t-2} \end{bmatrix} + \begin{bmatrix} \epsilon_{t+1} \\ 0 \\ 0 \end{bmatrix} \quad (19.19) \]

Even though this would not be the typical working representation of the AR(2) model, we show this form for transparency.

The Kalman filter addresses the problem of trying to estimate the state of a discrete time process as described here. The parameters of any model that can be written in state-space form can be estimated using the Kalman filter.

The Kalman Filter

A very general and extremely useful recursion algorithm is the Kalman filter (Kalman, 1960; Kalman & Bucy, 1961). It goes beyond the regression algorithms shown previously by allowing, among other things, time-varying parameters. Following Pullock (2003), we simply rewrite the least squares equations and provide names for some of the terms. This

ad hoc approach allows us to define the filter gain that represents the improvement of the estimate obtained by the updating algorithm. As we will see, all of the concepts we define represent simple combinations of the quantities we have already dealt with.

The covariance matrix of the regression parameters is \( C(\beta) = \sigma^2(X^T X)^{-1} \). If we define \( P_t = (X^T X)^{-1} \), then \( P(\beta) = \sigma^2 P_t \) and equation 19.14 becomes

\[ b_t = b_{t-1} + P_t^x (y_t - x_t b_{t-1}). \quad (19.20) \]

Now we give some definitions:

\[ h_t = y_t - X_t b_{t-1} \quad \text{prediction error} \]

\[ \sigma^2 / f_t = \sigma^2 (1 + (X_t - x_t) P_{t-1} x_t) \quad \text{dispersion error} \quad (19.21) \]

\[ \kappa_t = P_{t-1} x_t f_t^{-1} \quad \text{filter gain} \]

With these definitions we can express the recursion. For the parameter updating equation

\[ b_t = b_{t-1} + \kappa_t h_t \quad \text{parameter recursion} \quad (19.22) \]

the update is the prediction error multiplied by the filter gain, which can be thought of as a weight indicating the importance of the prediction error in determining the regression coefficient. For the dispersion updating equation, a similar equation can be derived:

\[ P_t = (I - \kappa_t x_t x_t^T) P_{t-1} \quad \text{dispersion recursion} \quad (19.23) \]

These two equations come directly from regression.

In the simplest sense the Kalman filter is a generalization of the previous recursion equations. The broad utility and incredible flexibility of the filter make it one of the most important technological innovations of the 20th century.

Given a state-space model in which observations arrive one occasion at a time, we can consider three different types of statistical estimation of the state: \( x(-) \) filtering, \( x(+) \) uses all of the data up to but not including \( x \); \( x(+) \) updating the prediction based on the latest observation; and \( x(0) \) smoothing. \( x(0) \) uses all of the observations to essentially filter to the end of the data set and work back to the newest observation.

The filter consists of a set of equations that first make a prediction of the next observation to arrive. Once those new data arrive, the prediction is adjusted based on the discrepancy between the prediction and the actual data point. The degree of adjustment is determined by the Kalman gain. The next observation arrives, and then the process is repeated.

This updating observation by observation allows some important additions to ordinary regression estimation in time series. One of the most important of these is the ability to have all parameters in the model time varying. This allows for possibilities in time series modeling well beyond many of the more standard techniques. Very readable discussion applications and variants of the basic filter appear in Grewal and Andrews (2001) and Simon (2006).
Conclusion

In this chapter we reviewed some of the issues regarding the estimation of person-specific and single-subject models. The possibility of estimating truly person-specific models becomes available only with the collection of time series data. At this point we remind the reader of the hopefully obvious fact that even though many methods are not truly person specific, those methods can still provide extremely helpful and useful information. However, when the data can be feasibly collected, in order to allow the possibility of different models for different individuals, time series person-specific approaches are certainly important to consider.

References

developmental processes on an individual level using non-stationary time series modeling. 


**CHAPTER 20**

Dynamic Factor Analysis and Control of Developmental Processes

**PETER C. M. MOLENAAR**

**LAWRENCE L. LO**

**Central Issues**

From a methodological perspective, developmental psychology always has been somewhat special in comparison with the other psychological subdisciplines, such as experimental psychology and social psychology. This is evident in the work, for instance, of Jean Piaget, one of the founding fathers of developmental psychology. Piaget's stage-wise developmental model of cognitive development was based on the *méthode clinique*, a special observational approach that gave rise to much discussion in the English-speaking scientific world. Brainard's (1978) critique of Piaget's stage-wise developmental model is well known and typical, classifying that model as being merely descriptive and proposing a Markovian learning model as an alternative. This long-standing discussion about the merits of Piaget's stage-wise developmental theory was resolved by Molenaar (1986a, 1986b) and van der Maas and Molenaar (1992), who showed, using a mathematical nonlinear dynamics approach, that developmental systems generally will display sudden qualitative changes in their evolution (so-called catastrophes, phase transitions, or bifurcations) of the kind described by Piaget. Consecutively, the catastrophe theoretical approach vindicated essential aspects of Piaget's model both in empirical research (e.g., Jansen & van der Maas, 2001, 2002; Rijmersma & Molenaar, 2004) and in theoretical respects (Wagenmakers, Molenaar, Grasmans, Hartelman, & van der Maas, 2005).

The importance of nonlinear dynamics approaches within psychology certainly is not confined to developmental psychology, but these approaches had an especially profound impact on developmental psychological methodology and theory, starting as early as the 1980s and lasting up to the present (cf. Spencer, Thomas, & McClelland, 2009). Similar