Description

To demonstrate a latent growth model, we use the following data:

```
. use http://www.stata-press.com/data/r13/sem_lcm
. describe
Contains data from http://www.stata-press.com/data/r13/sem_lcm.dta
    obs:     359
   vars:     4
  size:   5,744 (_dta has notes)

     variable name     storage  display       value label

     lncrime0  float     %9.0g      ln(crime rate) in Jan & Feb
     lncrime1  float     %9.0g      ln(crime rate) in Mar & Apr
     lncrime2  float     %9.0g      ln(crime rate) in May & Jun
     lncrime3  float     %9.0g      ln(crime rate) in Jul & Aug
```

Sorted by: notes

_dta:

See *Latent growth models* in [SEM] intro 5 for background.

Remarks and examples

Remarks are presented under the following headings:

- Fitting the model
- Fitting the model with the Builder
Fitting the model

We fit the following model:

```
. sem (lncrime0 <- Intercept@1 Slope@0 _cons@0)
> (lncrime1 <- Intercept@1 Slope@1 _cons@0)
> (lncrime2 <- Intercept@1 Slope@2 _cons@0)
> (lncrime3 <- Intercept@1 Slope@3 _cons@0),
> latent(Intercept Slope)
> var(e.lncrime0@var e.lncrime1@var
> e.lncrime2@var e.lncrime3@var)
> means(Intercept Slope)
```

Endogenous variables
Measurement: lncrime0 lncrime1 lncrime2 lncrime3
Exogenous variables
Latent: Intercept Slope

Fitting target model:
Iteration 0: log likelihood = -1034.1038
Iteration 1: log likelihood = -1033.9044
Iteration 2: log likelihood = -1033.9037
Iteration 3: log likelihood = -1033.9037

Structural equation model Number of obs = 359
Estimation method = ml
Log likelihood = -1033.9037

( 1) [Incrime0]Intercept = 1
( 2) [Incrime1]Intercept = 1
( 3) [Incrime1]Slope = 1
( 4) [Incrime2]Intercept = 1
( 5) [Incrime2]Slope = 2
( 6) [Incrime3]Intercept = 1
( 7) [Incrime3]Slope = 3
( 8) [var(e.Incrime0)]_cons - [var(e.Incrime3)]_cons = 0
( 9) [var(e.Incrime1)]_cons - [var(e.Incrime3)]_cons = 0
(10) [var(e.Incrime2)]_cons - [var(e.Incrime3)]_cons = 0
(11) [Incrime0]_cons = 0
(12) [Incrime1]_cons = 0
(13) [Incrime2]_cons = 0
(14) [Incrime3]_cons = 0
```
<table>
<thead>
<tr>
<th>Measurement</th>
<th>QIM</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>lncri-0 &lt;-</td>
<td>Coef.</td>
<td>Std. Err.</td>
<td>z</td>
<td>P&gt;</td>
<td>z</td>
</tr>
<tr>
<td>Intercept</td>
<td>1 (constrained)</td>
<td>0 (constrained)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>_cons</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>lncri-1 &lt;-</td>
<td>1 (constrained)</td>
<td>1 (constrained)</td>
<td>0 (constrained)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Slope</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>_cons</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>lncri-2 &lt;-</td>
<td>1 (constrained)</td>
<td>2 (constrained)</td>
<td>0 (constrained)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Slope</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>_cons</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>lncri-3 &lt;-</td>
<td>1 (constrained)</td>
<td>3 (constrained)</td>
<td>0 (constrained)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Slope</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>_cons</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>mean(Inter-</td>
<td>5.337915</td>
<td>.0407501</td>
<td>130.99</td>
<td>0.000</td>
<td>5.258047</td>
</tr>
<tr>
<td>t)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>mean(Slope)</td>
<td>.1426952</td>
<td>.0104574</td>
<td>13.65</td>
<td>0.000</td>
<td>.1221992</td>
</tr>
<tr>
<td>var(e.lncr-</td>
<td>.0981956</td>
<td>.0051826</td>
<td>.0885457</td>
<td>.1088972</td>
<td></td>
</tr>
<tr>
<td>0)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>var(e.lncr-</td>
<td>.0981956</td>
<td>.0051826</td>
<td>.0885457</td>
<td>.1088972</td>
<td></td>
</tr>
<tr>
<td>1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>var(e.lncr-</td>
<td>.0981956</td>
<td>.0051826</td>
<td>.0885457</td>
<td>.1088972</td>
<td></td>
</tr>
<tr>
<td>2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>var(e.lncr-</td>
<td>.0981956</td>
<td>.0051826</td>
<td>.0885457</td>
<td>.1088972</td>
<td></td>
</tr>
<tr>
<td>3)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>var(Inter-</td>
<td>.527409</td>
<td>.0446436</td>
<td>.4467822</td>
<td>.6225858</td>
<td></td>
</tr>
<tr>
<td>t)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>var(Slope)</td>
<td>.0196198</td>
<td>.0031082</td>
<td>.0143829</td>
<td>.0267635</td>
<td></td>
</tr>
<tr>
<td>cov(Inter-</td>
<td>-.034316</td>
<td>.0088848</td>
<td>-3.86</td>
<td>0.000</td>
<td>-.0517298</td>
</tr>
<tr>
<td>t, Slope)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes:

1. In this example, we have repeated measures of the crime rate in 1995. We will assume that the underlying rate grows linearly.

2. As explained in *Latent growth models* in [SEM] intro 5, we assume

\[ \text{lncrime}_i = \text{Intercept} + i \times \text{Slope} \]

3. *sem* does not usually report the means of latent exogenous variables because *sem* automatically includes the identifying constraint that the means are 0; see *How sem (gsem) solves the problem for you* in [SEM] intro 4 and see *Default normalization constraints* in [SEM] sem.

   In this case, *sem* did not constrain the means to be 0 because we specified *sem*’s means() option. In particular, we specified means(Intercept Slope), which said not to constrain the means of those two exogenous latent variables and to report the estimated result.

   Our model was identified even without the usual 0 constraints on Intercept and Slope because we specified enough other constraints.

4. We estimate the Intercept to have mean 5.34 and the mean Slope to be 0.14 per two months. Remember, we have measured crime rates as log base $e$ crime rates.
5. It might help some to think of this as a mixed model:

   . generate id = _n
   . reshape long lncrime, i(id) j(year)
   . mixed lncrime year || id:year, cov(unstructured) mle var

   The mean Intercept and Slope are what mixed would refer to as the coefficients in the
   fixed-effects part of the model.

   Acoc (2013, chap. 4) discusses the use of sem to fit latent growth-curve models in more detail. 
   Acoc demonstrates extensions to the basic model we fit here, such as including time-varying and 
   time-invariant covariates in the model.

**Fitting the model with the Builder**

   Use the diagram above for reference.

   1. Open the dataset.
      In the Command window, type
      
      . use http://www.stata-press.com/data/r13/sem_lcm

   2. Open a new Builder diagram.
      Select menu item **Statistics > SEM (structural equation modeling) > Model building and estimation**.

   3. Create the four repeated measurements for log crime rate.
      Select the Add Observed Variables Set tool, and then click in the diagram about one-fourth
      of the way in from the left and about one-third of the way up from the bottom.
      In the resulting dialog box,
      a. select the **Select variables** radio button (it may already be selected);
      b. use the **Variables** control to select the variables in this order: lncrime0, lncrime1, 
         lncrime2, and lncrime3;
      c. select **Horizontal** in the **Orientation** control;
      d. select the **Distances** tab;
      e. select .25 (inch) from the **Distance between variables** control;
      f. click on **OK**.
      If you wish, move the set of variables by clicking on any variable and dragging it.

   4. Create the two latent variables.
      Select the Add Latent Variables Set tool, and then click in the diagram about one-third of
      the way down from the top and above the left side of the lncrime0 rectangle.
      In the resulting dialog box,
      a. select the **Select variables** radio button (it may already be selected);
      b. type the latent variable names **Intercept** and **Slope** into the **Variables** control;
      c. select **Horizontal** in the **Orientation** control;
      d. select the **Distances** tab;
e. select 1 (inch) from the Distance between variables control;
f. click on OK.

If you wish, move the set of variables by clicking on any variable and dragging it.

We could have forgone specifying the distances between variables and simply dragged each variable where we wished after it was created.

5. Create the factor-loading paths.

a. Select the Add Path tool, 

b. Click in the bottom-left quadrant of the Intercept oval (it will highlight when you hover over it), and drag a path to the top of the lncrime0 rectangle (it will highlight when you can release to connect the path).

c. Continuing with the tool, create the following paths by clicking first in the bottom of the latent variable and dragging it to the top of the observed (measurement) variable:

   Intercept -> lncrime1
   Intercept -> lncrime2
   Intercept -> lncrime3
   Slope -> lncrime0
   Slope -> lncrime1
   Slope -> lncrime2
   Slope -> lncrime3

6. Clean up the direction of the errors.

We want all the errors to be below the measurement variables.

a. Choose the Select tool, 

b. Click on the rectangle of any measurement variable whose associated error is not below it.

c. Click on one of the Error Rotation buttons, , in the Contextual Toolbar until the error is below the measurement variable.

Repeat this for all errors that are not below the measurement variables.

7. Create the covariance between Intercept and Slope.

a. Select the Add Covariance tool, 

b. Click in the top-right quadrant of the Intercept oval, and drag a covariance to the top left of the Slope oval.

8. Clean up paths and covariance.

If you do not like where a path has been connected to its variables, use the Select tool, to click on the path, and then simply click on where it connects to a rectangle and drag the endpoint. Similarly, you can change where the covariance connects to the latent variables by clicking on the covariance and dragging the endpoint. You can also change the bow of the covariance by clicking on the covariance and dragging the control point that extends from one end of the selected covariance.

9. Constrain the intercepts of the measurements to 0.

a. Choose the Select tool, .
b. Click on the rectangle for lncrime0. In the Contextual Toolbar, type 0 in the box and press Enter.

c. Repeat this process to add the 0 constraint on the intercept for lncrime1, lncrime2, and lncrime3.

10. Set constraints on the paths from Intercept to the measurements.

   a. Continue with the Select tool, 

   b. Click on the path from Intercept to lncrime0. In the Contextual Toolbar, type 1 in the box and press Enter.

   c. Repeat this process to add the 1 constraint on the following paths:

   \[
   \text{Intercept} \rightarrow \text{lncrime1} \\
   \text{Intercept} \rightarrow \text{lncrime2} \\
   \text{Intercept} \rightarrow \text{lncrime3}
   \]

11. Set constraints on the paths from Slope to the measurements.

   a. Continue with the Select tool, 

   b. Click on the path from Slope to lncrime0. In the Contextual Toolbar, type 0 in the box and press Enter.

   c. Click on the path from Slope to lncrime1. In the Contextual Toolbar, type 1 in the box and press Enter.

   d. Click on the path from Slope to lncrime2. In the Contextual Toolbar, type 2 in the box and press Enter.

   e. Click on the path from Slope to lncrime3. In the Contextual Toolbar, type 3 in the box and press Enter.

12. Set equality constraints on the error variances.

   a. Continue with the Select tool, 

   b. Click in the $\epsilon_1$ circle, which is the error term for lncrime0. In the Contextual Toolbar, type var in the box and press Enter.

   c. Repeat this process to add the var constraint on the three remaining error variances: $\epsilon_2$, $\epsilon_3$, and $\epsilon_4$.

13. Clean up placement of the constraints.

   From the SEM Builder menu, select Settings > Connections > Paths....

   In the resulting dialog box, do the following:

   a. Click on the Results tab.

   b. Click on the Result 1... button at the bottom left.

   c. In the Appearance of result 1 - paths dialog box that opens, choose 20 (%) in the Distance between nodes control.

   d. Click on OK on the Appearance of result 1 - paths dialog box.

   e. Click on OK on the Connection settings - paths dialog box.
14. Specify that the means of Intercept and Slope are to be estimated.
   a. Choose the Select tool, 🔄.
   b. Double-click on the Intercept oval.
   c. In the resulting dialog box, check the Estimate mean box.
   d. Click on OK.
   e. In the same manner, double-click on Slope and check the Estimate mean box.

15. Estimate.
   Click on the Estimate button, 🔄, in the Standard Toolbar, and then click on OK in the resulting SEM estimation options dialog box.

   You can open a completed diagram in the Builder by typing
   .wgetsem sem_lcm

Reference

Also see
[SEM] sem — Structural equation model estimation command