Technical tips on time series with Stata

Gustavo Sánchez
Senior Statistician
StataCorp

2011 Mexican Stata Users Group Meeting 2011
Outline

- Tip 1: Specifying the time structure
  - tsset
  - Date formats

- Tip 2: Why some predictions with -arima- do not match my manual computations - Kalman Filter recursions

- Tip 3: What is the initial shock for impulse response functions after -var-

- Tip 4: How do I fit my unobserved component model with -sspace-
  - linear regression and random walk

- Tip 5: How do I specify restrictions on the long-run cointegrating relationship in the VEC model
TIP 1: Specifying the time structure

- `tsset timevar [, options]`
  - Date frequency (daily, weekly, monthly,...)
  - Clocktime (hours, minutes, seconds,..., milliseconds)
  - Generic
  - `delta()`
  
  Example:
  ```
  tsset timevar, daily delta(7)
  lags in terms of seven days
  ```
TIP 1: Specifying the time structure

- Date formats
  - Example – Daily format

```stata
clear
input str12 date
   "1/01/2008"
   "1/02/2008"
   "1/03/2008"
   "1/04/2008"
   "1/05/2008"
end

generate mydate1=date(date,"DMY")
format mydate1 %td

generate mydate2=date(date,"DMY")
format mydate2 %tdmon-DD,_CCYY
```
TIP 1: Specifying the time structure

- Date formats
  - Example – Daily format

  . list date mydate1 mydate2

<table>
<thead>
<tr>
<th></th>
<th>date</th>
<th>mydate1</th>
<th>mydate2</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>1/01/2008</td>
<td>01jan2008</td>
<td>jan-01, 2008</td>
</tr>
<tr>
<td>2.</td>
<td>1/02/2008</td>
<td>01feb2008</td>
<td>feb-01, 2008</td>
</tr>
<tr>
<td>3.</td>
<td>1/03/2008</td>
<td>01mar2008</td>
<td>mar-01, 2008</td>
</tr>
<tr>
<td>4.</td>
<td>1/04/2008</td>
<td>01apr2008</td>
<td>apr-01, 2008</td>
</tr>
<tr>
<td>5.</td>
<td>1/05/2008</td>
<td>01may2008</td>
<td>may-01, 2008</td>
</tr>
</tbody>
</table>
TIP 1: Specifying the time structure

- **Date formats**
  - **Example – Daily format**

  ```
  . tsset mydate1
  time variable: mydate1, 01jan2008 to 01may2008, but with gaps
  delta: 1 day
  
  . list mydate1 if tin(01feb2008,01apr2008)
  
  +------------------------+
<table>
<thead>
<tr>
<th>mydate1</th>
</tr>
</thead>
</table>
  2. | 01feb2008 |
  3. | 01mar2008 |
  4. | 01apr2008 |
  +------------------------+
  ```
TIP 1: Specifying the time structure

- Date formats
  - Example – Clock format

```
clear
Input          str20    etime          y
             "06feb2010 12:40:00"  2
             "06feb2010 12:42:00"  5
             "06feb2010 12:44:00"  7
             "06feb2010 12:46:00"  6
             "06feb2010 12:48:00"  9
end

generate double mytime = clock(etime, "DMY hms")
format mytime %tc DMYHH:MM:SS
```
TIP 1: Specifying the time structure

- **Date formats**
  - Example – Clock format
    ```
    . tsset mytime, delta(2 minute)
    time variable: mytime, 06feb2010 12:40:00 to 06feb2010 12:48:00
    delta: 2 minutes
    
    . generate my_ly=l.y
    (1 missing value generated)
    
    . list mytime y ly my_ly
      +---------------------------------+--
      |             mytime   y    my_ly |
      |---------------------------------|
      | 1. | 06feb2010 12:40:00   2        . |
      | 2. | 06feb2010 12:42:00   5        2 |
      | 3. | 06feb2010 12:44:00   7        5 |
      | 4. | 06feb2010 12:46:00   6        7 |
      | 5. | 06feb2010 12:48:00   9        6 |
    +---------------------------------+
    ```
TIP 2: Predictions with -arima- Kalman Filter recursions

- Let’s consider the following moving average (MA1) model:

\[ y_t = \alpha + \theta \epsilon_{t-1} + \epsilon_t ; \quad \epsilon_t \sim i.i.d. N(0, \sigma^2) \]

Command line to fit the model:

```
arima y, ma(1)
```

And we get the predictions with:

```
predict double y_hat
```
TIP 2: Predictions with -arima- Kalman Filter recursions

- Users try to manually reproduce the predictions with:

\[ \hat{Y}_t = \hat{\alpha} + \hat{\theta} \times \hat{\epsilon}_{t-1} \]
\[ \hat{\epsilon}_{t-1} = Y_{t-1} - \hat{Y}_{t-1} \]

- However, the results do not match the predictions obtained with:

```
predict double y_hat
```

**WHY?**
TIP 2: Predictions with -arima- Kalman Filter recursions

- Code for manual predictions

```stata
use http://www.stata-press.com/data/r11/lutkepohl,clear
arima dlinvestment, ma(1)
predict double yhat
scalar b0 = _b[_cons]
scalar t1 = [ARMA]_b[L1.ma]
gen double my_yhat = b0
gen double myehat = dlinvestment - b0 in 2
forvalues i = 3/91 {
    qui replace my_yhat = my_yhat ///
        + t1*L.myehat in `i'
    qui replace myehat = dlinvestment - my_yhat in `i'
}
```
TIP 2: Predictions with -arima-

Kalman Filter recursions

- List first 12 predictions

```
.list qtr yhat my_yhat in 1/13, sep(11)
```

<table>
<thead>
<tr>
<th>qtr</th>
<th>yhat</th>
<th>my_yhat</th>
</tr>
</thead>
<tbody>
<tr>
<td>1960q1</td>
<td>.01686688</td>
<td>.01686688</td>
</tr>
<tr>
<td>1960q2</td>
<td>.01686688</td>
<td>.01686688</td>
</tr>
<tr>
<td>1960q3</td>
<td>.02052151</td>
<td>.02062398</td>
</tr>
<tr>
<td>1960q4</td>
<td>.01478403</td>
<td>.0147996</td>
</tr>
<tr>
<td>1961q1</td>
<td>.01312365</td>
<td>.01312617</td>
</tr>
<tr>
<td>1961q2</td>
<td>.00326376</td>
<td>.00326418</td>
</tr>
<tr>
<td>1961q3</td>
<td>.02471242</td>
<td>.02471249</td>
</tr>
<tr>
<td>1961q4</td>
<td>.01691061</td>
<td>.01691062</td>
</tr>
<tr>
<td>1962q1</td>
<td>.01412974</td>
<td>.01412975</td>
</tr>
<tr>
<td>1962q2</td>
<td>.00643301</td>
<td>.00643301</td>
</tr>
<tr>
<td>1962q3</td>
<td>.01940009</td>
<td>.0194001</td>
</tr>
<tr>
<td>1962q4</td>
<td>.01649863</td>
<td>.01649863</td>
</tr>
<tr>
<td>1963q1</td>
<td>.01749646</td>
<td>.01749646</td>
</tr>
</tbody>
</table>
TIP 2: Predictions with -arima- Kalman Filter recursions

- Stata uses the recursive formula for the Kalman filter prediction based on:

\[
\hat{\epsilon}_{t-1} = \left[ \frac{\hat{\sigma}^2}{\hat{\sigma}^2 + \hat{\theta}^2 \times p_{t-1}} \right] \times (Y_{t-1} - \hat{Y}_{t-1}) \quad \hat{\epsilon}_0 = 0
\]

- Where:

\[
p_t = \left[ \frac{\hat{\sigma}^2 \times \hat{\theta}^2 \times p_{t-1}}{\hat{\sigma}^2 + \hat{\theta}^2 \times p_{t-1}} \right] \quad p_1 = \hat{\sigma}^2
\]

\(\hat{\sigma}^2\) estimated variance of the white noise disturbance
TIP 2: Predictions with -arima- Kalman Filter recursions

use http://www.stata-press.com/data/r11/lutkepohl, clear
arima dlinvestment, ma(1)
predict double yhat

** Coefficient estimates and sigma^2 from ereturn list **
scalar b0 = _b[_cons]
scalar t1 = [ARMA]_b[L1.ma]
scalar sigma2 = e(sigma)^2

** pt and shrinking factor for the first two observations**
gen double pt= sigma2 in 1/2
gen double myratio=(sigma2)/(sigma2+t1^2*pt) in 2

** Predicted series and errors for the first two observations **
gen double my_yhat = b0
generate double myehat = myratio*(dlinvestment - my_yhat) in 2

** Predictions with the Kalman filter recursions **
forvalues i = 3/91 {
    qui replace my_yhat = my_yhat + t1*l.myehat in `i'
    qui replace pt= (sigma2)*(t1^2)*(L.pt)/(sigma2+t1^2*L.pt) in `i'
    qui replace myratio=(sigma2)/(sigma2+t1^2*pt) in `i'
    qui replace myehat=myratio*(dlinvestment - my_yhat) in `i'
}

TIP 2: Predictions with -arima- Kalman Filter recursions

- List first 10 predictions

```
.list qtr yhat my_yhat pt myratio in 1/10
```

<table>
<thead>
<tr>
<th>qtr</th>
<th>yhat</th>
<th>my_yhat</th>
<th>pt</th>
<th>myratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>1960q1</td>
<td>0.01686688</td>
<td>0.01686688</td>
<td>0.00192542</td>
<td>.</td>
</tr>
<tr>
<td>1960q2</td>
<td>0.01686688</td>
<td>0.01686688</td>
<td>0.00192542</td>
<td>0.97272668</td>
</tr>
<tr>
<td>1960q3</td>
<td>0.02052151</td>
<td>0.02052151</td>
<td>0.0005251</td>
<td>0.99923589</td>
</tr>
<tr>
<td>1960q4</td>
<td>0.01478403</td>
<td>0.01478403</td>
<td>1.471e-06</td>
<td>0.99997858</td>
</tr>
<tr>
<td>1961q1</td>
<td>0.01312365</td>
<td>0.01312365</td>
<td>4.125e-08</td>
<td>0.9999994</td>
</tr>
<tr>
<td>1961q2</td>
<td>0.00326376</td>
<td>0.00326376</td>
<td>1.157e-09</td>
<td>0.9999998</td>
</tr>
<tr>
<td>1961q3</td>
<td>0.02471242</td>
<td>0.02471242</td>
<td>3.243e-11</td>
<td>1</td>
</tr>
<tr>
<td>1961q4</td>
<td>0.01691061</td>
<td>0.01691061</td>
<td>9.092e-13</td>
<td>1</td>
</tr>
<tr>
<td>1962q1</td>
<td>0.01412974</td>
<td>0.01412974</td>
<td>2.549e-14</td>
<td>1</td>
</tr>
<tr>
<td>1962q2</td>
<td>0.00643301</td>
<td>0.00643301</td>
<td>7.147e-16</td>
<td>1</td>
</tr>
</tbody>
</table>
TIP 3: Initial shock for Impulse response functions (IRF) after -var- VAR model

$$\Delta Y_t = \alpha + \beta_1 \Delta Y_{t-1} + \beta_2 \Delta Y_{t-2} + \ldots + \beta_p \Delta Y_{t-p} + \varepsilon_t$$

Where:

$$Y_t = (Y_{1t}, Y_{2t}, \ldots, Y_{kt}) : \text{l(1) Endogenous variables}$$

$$\beta_i : \text{Matrix with coefficients associated to lag } i$$

$$\alpha : \text{Vectors with coefficients associated to the intercepts}$$

$$\varepsilon_t : \text{Vector with innovations}$$
TIP 3: Initial shock for Impulse response functions (IRF) after *-var-

- Orthogonalized IRF functions for a shock in Y1

Graphs by irfname, impulse variable, and response variable
TIP 3: Initial shock for Impulse response functions after `-var-`

What is the magnitude of the shock in the IRF graph?

- `-irf graph, irf-`: simple IRF
  - correspond to one-time unit increase
  - the effects do not have a causal interpretation
TIP 3: Initial shock for Impulse response functions after -var-

What is the magnitude of the shock in the IRF graph?

- **-irf graph,oirf-**: orthogonal IRF
  - orthogonalization is produced via the Cholesky decomposition
  - the magnitude of the shock corresponds to one unit standard deviation

- **-irf graph,sirf-** structural IRF
  - -irf graph,sirf- IRF functions are derived from the constraints imposed on the SVAR
  - the magnitude of the shock corresponds to one unit standard deviation

Graphs by irfname, impulse variable, and response variable
TIP 3: Initial shock for Impulse response functions after -var-

- Let’s fit a VAR model:

  use http://www.stata-press.com/data/r11/lutkepohl
  var dlinvestment dlincome, lags(1/2) dfk
TIP 3: Initial shock for Impulse response functions after -var-

```
. var dlinvestment dlincome, lags(1/2) dfk
```

Vector autoregression

<table>
<thead>
<tr>
<th>Equation</th>
<th>Parms</th>
<th>RMSE</th>
<th>R-sq</th>
<th>chi2</th>
<th>P&gt;chi2</th>
</tr>
</thead>
<tbody>
<tr>
<td>dlinvestment</td>
<td>5</td>
<td>.04424</td>
<td>0.0856</td>
<td>8.32989</td>
<td>0.0802</td>
</tr>
<tr>
<td>dlincome</td>
<td>5</td>
<td>.011403</td>
<td>0.1027</td>
<td>10.1916</td>
<td>0.0373</td>
</tr>
</tbody>
</table>

|              | Coef.   | Std. Err. | z     | P>|z|    | [95% Conf. Interval] |
|--------------|---------|------------|-------|--------|----------------------|
| dlinvestment |         |            |       |        |                      |
| L1. | -.2274192 | .1053092 | -2.16 | 0.031  | -.4338214 - .021017  |
| L2. | -.1159636 | .1057698 | -1.10 | 0.273  | -.3232686 .0913415   |
| dlincome    |         |            |       |        |                      |
| L1. | .7103053  | .3948248  | 1.80  | 0.072  | -.0635372 1.484148   |
| L2. | .5149489  | .3935121  | 1.31  | 0.191  | -.2563206 1.286218   |
| _cons      | -.0012273 | .0111362 | -0.11 | 0.912  | -.0230539 .0205993   |

<p>| | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>dlincome</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L1.</td>
<td>.0597466</td>
<td>.0271441</td>
<td>2.20</td>
<td>0.028</td>
<td>.0065451  .1129481</td>
</tr>
<tr>
<td>L2.</td>
<td>.0563513</td>
<td>.0272629</td>
<td>2.07</td>
<td>0.039</td>
<td>.002917 .1097855</td>
</tr>
<tr>
<td>dlincome</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L1.</td>
<td>.0209461</td>
<td>.1017687</td>
<td>0.21</td>
<td>0.837</td>
<td>-.1785169 .220409</td>
</tr>
<tr>
<td>L2.</td>
<td>.0833252</td>
<td>.1014303</td>
<td>0.82</td>
<td>0.411</td>
<td>-.1154745 .2821249</td>
</tr>
<tr>
<td>_cons</td>
<td>.0150368</td>
<td>.0028704</td>
<td>5.24</td>
<td>0.000</td>
<td>.0094108 .0206627</td>
</tr>
</tbody>
</table>
TIP 3: Initial shock for Impulse response functions after -var-

- Plot the IRF function for a shock in dlinvestment

```stata
use http://www.stata-press.com/data/r11/lutkepohl
var dlinvestment dlincome, lags(1/2) dfk
irf create order1, step(10) set(myirf1,replace)
irf graph oirf, impulse(dlinvestment) ///
response(dlinvestment dlincome)
```
TIP 3: Initial shock for Impulse response functions after -var-

- Plot the IRF function for a shock in dlinvestment
  \texttt{irf graph oirf, impulse(dlinvestment) /// response(dlinvestment dlincome)}

Graphs by irfname, impulse variable, and response variable

Graphs by irfname, impulse variable, and response variable

\texttt{orthogonalized irf}
TIP 3: Initial shock for Impulse response functions after -var-

Table for the OIRF function for a shock in dlinvestment

```
. irf table oirf, irf(order1) impulse(dlinvestment) ///
    response(dlincome dlinvestment) ///
Results from order1

<table>
<thead>
<tr>
<th>step</th>
<th>(1)</th>
<th>(1)</th>
<th>(1)</th>
<th>(2)</th>
<th>(2)</th>
<th>(2)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>oirf</td>
<td>Lower</td>
<td>Upper</td>
<td>oirf</td>
<td>Lower</td>
<td>Upper</td>
</tr>
<tr>
<td>0</td>
<td>.001641</td>
<td>-.000715</td>
<td>.003998</td>
<td>.04424</td>
<td>.037741</td>
<td>.050739</td>
</tr>
<tr>
<td>1</td>
<td>.002678</td>
<td>.000241</td>
<td>.005114</td>
<td>-.008895</td>
<td>-.018459</td>
<td>.000669</td>
</tr>
<tr>
<td>2</td>
<td>.002154</td>
<td>-.000283</td>
<td>.004592</td>
<td>-.00036</td>
<td>-.009879</td>
<td>.009159</td>
</tr>
<tr>
<td>3</td>
<td>-.000255</td>
<td>-.001358</td>
<td>.000849</td>
<td>.004022</td>
<td>.001068</td>
<td>.009113</td>
</tr>
<tr>
<td>4</td>
<td>.000394</td>
<td>-.000347</td>
<td>.001136</td>
<td>.000056</td>
<td>-.002258</td>
<td>.000237</td>
</tr>
<tr>
<td>5</td>
<td>.000217</td>
<td>-.000207</td>
<td>.000641</td>
<td>-.00033</td>
<td>-.002245</td>
<td>.001585</td>
</tr>
<tr>
<td>6</td>
<td>.000021</td>
<td>-.000237</td>
<td>.000279</td>
<td>.000426</td>
<td>-.000457</td>
<td>.001309</td>
</tr>
<tr>
<td>7</td>
<td>.000025</td>
<td>-.000101</td>
<td>.000152</td>
<td>.000068</td>
<td>-.000353</td>
<td>.000488</td>
</tr>
<tr>
<td>8</td>
<td>.00003</td>
<td>-.000055</td>
<td>.000116</td>
<td>-.000036</td>
<td>-.000356</td>
<td>.000284</td>
</tr>
<tr>
<td>9</td>
<td>4.4e-06</td>
<td>-.000034</td>
<td>.000043</td>
<td>.000035</td>
<td>-.000074</td>
<td>.000143</td>
</tr>
<tr>
<td>10</td>
<td>2.7e-06</td>
<td>-.000022</td>
<td>.000027</td>
<td>.000015</td>
<td>-.000063</td>
<td>.000093</td>
</tr>
</tbody>
</table>
```

95% lower and upper bounds reported

(1) irfname = order1, impulse = dlinvestment, and response = dlincome
(2) irfname = order1, impulse = dlinvestment, and response = dlinvestment
**TIP 3: Initial shock for Impulse response functions after -var-**

- Table for the IRF function for a shock in dlinvestment

```
. irf table irf, irf(order1) impulse(dlinvestment) ///
  >     response(dlincome dlinvestment)

Results from order1

<table>
<thead>
<tr>
<th>step</th>
<th>(1) irf</th>
<th>(1) Lower</th>
<th>(1) Upper</th>
<th>(2) irf</th>
<th>(2) Lower</th>
<th>(2) Upper</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>.059747</td>
<td>.004985</td>
<td>.114509</td>
<td>-.227419</td>
<td>-.439876</td>
<td>-.014963</td>
</tr>
<tr>
<td>2</td>
<td>.044015</td>
<td>-.010388</td>
<td>.098419</td>
<td>-.021806</td>
<td>-.237257</td>
<td>.193646</td>
</tr>
<tr>
<td>3</td>
<td>-.008218</td>
<td>-.032283</td>
<td>.015847</td>
<td>.093362</td>
<td>-.027102</td>
<td>.213826</td>
</tr>
<tr>
<td>4</td>
<td>.007845</td>
<td>-.007056</td>
<td>.022745</td>
<td>-.001875</td>
<td>-.054015</td>
<td>.050264</td>
</tr>
<tr>
<td>5</td>
<td>.004629</td>
<td>-.004709</td>
<td>.013967</td>
<td>-.00906</td>
<td>-.054602</td>
<td>.036483</td>
</tr>
<tr>
<td>6</td>
<td>.000104</td>
<td>-.005125</td>
<td>.005332</td>
<td>.009605</td>
<td>-.010735</td>
<td>.029945</td>
</tr>
<tr>
<td>7</td>
<td>.000451</td>
<td>-.002119</td>
<td>.003022</td>
<td>.001323</td>
<td>-.00833</td>
<td>.010977</td>
</tr>
<tr>
<td>8</td>
<td>.000638</td>
<td>-.001136</td>
<td>.002413</td>
<td>-.001041</td>
<td>-.008544</td>
<td>.006462</td>
</tr>
<tr>
<td>9</td>
<td>.000063</td>
<td>-.000688</td>
<td>.000814</td>
<td>.000769</td>
<td>-.001641</td>
<td>.003179</td>
</tr>
<tr>
<td>10</td>
<td>.000042</td>
<td>-.000454</td>
<td>.000538</td>
<td>.00032</td>
<td>-.001466</td>
<td>.002105</td>
</tr>
</tbody>
</table>
```

95% lower and upper bounds reported

(1) irfname = order1, impulse = dlinvestment, and response = dlincome
(2) irfname = order1, impulse = dlinvestment, and response = dlinvestment
TIP 4: How do I fit my unobserved component model with \texttt{sspace}-

- State Space representation

\[
\begin{align*}
    z_t &= A z_{t-1} + B x_t + C \varepsilon_t \\
    y_t &= D z_t + F w_t + G \nu_t
\end{align*}
\]

Where:

- \( z_t \): is an \( m \times 1 \) vector of unobserved state variables;
- \( x_t \): is a \( k_x \times 1 \) vector of exogenous variables;
- \( \varepsilon_t \): is a \( q \times 1 \) vector of state-error terms, \((q \leq m)\);
- \( y_t \): is an \( n \times 1 \) vector of observed endogenous variables;
- \( w_t \): is a \( k_w \times 1 \) vector of exogenous variables;
- \( \nu_t \): is an \( r \times 1 \) vector of observation-error terms, \((r \leq n)\); and
- A, B, C, D, F, and G are parameter matrices.
TIP 4: How do I fit my unobserved component model with —sspace—

- State Space representation for linear regression

\[
\begin{align*}
    z_t &= \varepsilon_t \\
    y_t &= z_t + \alpha + \beta w_t
\end{align*}
\]

Command specification

- constraint 1 \([z]L.z = 0\)
- constraint 2 \([y]z = 1\)
- sspace (z L.z, state noconstant) ///
  (y w z,noerror ), constraints(1/2)
TIP 4:

State Space estimation for linear regression

use http://www.stata-press.com/data/r11/lutkepohl,clear
constraint 1 [z]L.z = 0
constraint 2 [dlinvestment]z = 1
sspace (z L.z, state noconstant) (dlinvestment dlincome z,noerror ), constraints(1/2) nolog

State-space model
Sample: 1960q2 - 1982q4                           Number of obs   =         91
Wald chi2(1)    =       0.88                           Prob > chi2     =     0.3487
Log likelihood =  154.44197                        Number of obs   =         91

------------------------------------------------------------------------------
dlinvestment |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+--------------------------------------------------
z |          1          .        .       .            .           .
dlincome |   .3668678   .3914794     0.94   0.349   -.4004178    1.134153
_cons |   .0096556    .008925     1.08   0.279   -.007837    .0271483
-------------+--------------------------------------------------
var(z)       |   .0019651   .0002913     6.75   0.000     .0013941    .0025361
-------------+--------------------------------------------------

Note: Tests of variances against zero are conservative and are provided only for reference.
TIP 4: How do I fit my unobserved component model with –sspace–

- Random Walk

\[ y_t = y_{t-1} + \varepsilon \]

- State Space representation

\[ z_t = z_{t-1} + \varepsilon_t \]
\[ y_t = z_t \]

- Command specification

constraint 1 \([z]L.z = 1\)
constraint 2 \([y]z = 1\)
sspace \((z L.z, \text{state noconstant})\)
\((y \ z, \text{noerror noconstant}), \text{constraints(1/2)}\)
TIP 4:

- **State Space estimation for Random Walk**

```
constraint 1 [z]L_z = 1
constraint 2 [dlinvestment]z = 1
sspace (z L.z, state noconstant) (dlinvestment z, noerror noconstant), constraints(1/2) nolog
```

**State-space model**

Sample: 1960q2 - 1982q4                           Number of obs   =    91
Log likelihood =  112.76541

```
                      OIM
            dlinvestment |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
            z  |          1          .        .       .            .           .
            z  |          1          .        .       .            .           .
-------------+----------------------------------------------------------------
  dlinvestment |          1          .        .       .            .           .
            z  |          1          .        .       .            .           .
-------------+----------------------------------------------------------------
  var(z)  |   .0046812   .0006978     6.71   0.000     .0033135     .006049
-------------+----------------------------------------------------------------
```

Note: Model is not stationary.
Note: Tests of variances against zero are conservative and are provided only for reference.
TIP 5: VEC – Johansen identification

Reduced form for a VEC model

\[ \Delta z_t = a + bt - \alpha \beta z_{t-1} + \sum_{i=1}^{p-1} \Gamma_i \Delta z_{t-i} + \nu_t \]

Where:

\( z_t \) : I(1) Endogenous variables

\( \alpha \beta \) : Matrices containing the long-run adjustment coefficients and coefficients for the cointegrating relationships

\( \Gamma_i \) : Matrix with coefficients associated to short-run dynamic effects

\( a, b \) : Vectors with coefficients associated to the intercepts and trends

\( \nu_t \) : Vector with innovations
TIP 5:

Example: VEC with three endogenous variables

\[
\begin{bmatrix}
\Delta z_{1t} \\
\Delta z_{2t} \\
\Delta z_{3t}
\end{bmatrix} = \alpha + bt - \alpha \beta
\begin{bmatrix}
z_{1t-1} \\
z_{2t-1} \\
z_{3t-1}
\end{bmatrix} + \sum_{i=1}^{p-1} \Gamma_i
\begin{bmatrix}
\Delta z_{1t-i} \\
\Delta z_{2t-i} \\
\Delta z_{3t-i}
\end{bmatrix} + \nu_t
\]

Where:

\[
\alpha \beta z_{t-1} = \begin{bmatrix}
\alpha_{11} & \alpha_{12} \\
\alpha_{21} & \alpha_{22} \\
\alpha_{31} & \alpha_{32}
\end{bmatrix}
\begin{bmatrix}
\beta_{11} & \beta_{12} & \beta_{13} \\
\beta_{21} & \beta_{22} & \beta_{23} \\
\end{bmatrix}
\begin{bmatrix}
z_{1t-1} \\
z_{2t-1} \\
z_{3t-1}
\end{bmatrix}
\]

- Identifying \( \alpha \) and \( \beta \) requires \( r^2 \) restrictions (\( r \): number of cointegrating vectors).
- Johansen FIML estimation identifies \( \alpha \) and \( \beta \) by imposing \( r^2 \) atheoretical restrictions.

\[
\alpha \beta z_{t-1} = \begin{bmatrix}
\alpha_{11} & \alpha_{12} \\
\alpha_{21} & \alpha_{22} \\
\alpha_{31} & \alpha_{32}
\end{bmatrix}
\begin{bmatrix}
1 & 0 & \beta_{13} \\
0 & 1 & \beta_{23} \\
\end{bmatrix}
\begin{bmatrix}
z_{1t-1} \\
z_{2t-1} \\
z_{3t-1}
\end{bmatrix}
\]
**TIP 5:**

- Restrictions based on Johansen normalization (Default)

use http://www.stata-press.com/data/r11/lutkepohl,clear
vec linvestment lincome lconsumption, rank(2) lags(2) noetable trend(none)

Vector error-correction model

Identification: beta is exactly identified

|       | Coef. | Std. Err. | z    | P>|z| | [95% Conf. Interval] |
|-------|-------|-----------|------|-------|----------------------|
| _ce1  |       |           |      |       |                      |
| linvestment | 1     |           |      |       |                      |
| lincome   | (omitted) |         |      |       |                      |
| lconsumption | -0.7943718 | 0.0125908 | -63.09 | 0.000 | -0.8190493 -0.7696942 |
| _ce2  |       |           |      |       |                      |
| linvestment | (omitted) |         |      |       |                      |
| lincome   | 1     |           |      |       |                      |
| lconsumption | -1.013321 | 0.0013846 | -731.87 | 0.000 | -1.016035 -1.010608 |

Johansen normalization restrictions imposed
TIP 5:

Instead of the Johansen atheoretical restrictions we could use economic theory to impose restrictions to identify $\alpha\beta$.

For example, let’s assume the following cointegrating equations:

\[
\begin{align*}
z_{1t-1} &= 0.75z_{2t-1} + \beta_{13}z_{3t-1} \\
z_{3t-1} &= .85z_{2t-1} + \beta_{21}z_{1t-1}
\end{align*}
\]

Which implies

\[
\alpha\beta z_{t-1} = \begin{bmatrix}
\alpha_{11} & \alpha_{12} \\
\alpha_{21} & \alpha_{22} \\
\alpha_{31} & \alpha_{32}
\end{bmatrix} \begin{bmatrix}
1 & -0.75 & \beta_{13} \\
\beta_{21} & -0.85 & 1
\end{bmatrix} \begin{bmatrix}
z_{1t-1} \\
z_{2t-1} \\
z_{3t-1}
\end{bmatrix}
\]
TIP 5:
- Restrictions specified by the user

<table>
<thead>
<tr>
<th>Constraint</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>[_ce1\text{linvestment}=1]</td>
</tr>
<tr>
<td>2</td>
<td>[_ce1\text{lincome}=-.75]</td>
</tr>
<tr>
<td>3</td>
<td>[_ce2\text{lconsumption}=1]</td>
</tr>
<tr>
<td>4</td>
<td>[_ce2\text{lincome}=-.85]</td>
</tr>
</tbody>
</table>

vec \text{linvestment lincome lconsumption}, rank(2) lags(2) noetable trend(none) bconstraints(1/4)

Identification: \text{beta} is exactly identified

( 1) \[\_ce1\text{linvestment} = 1\]
( 2) \[\_ce1\text{lincome} = -.75\]
( 3) \[\_ce2\text{lconsumption} = 1\]
( 4) \[\_ce2\text{lincome} = -.85\]

| beta | Coef.    | Std. Err. | z     | P>|z| | \[95\% Conf. Interval\] |
|------|----------|-----------|-------|------|-------------------------|
| _ce1 |          |           |       |      |                         |
| _ce1 | linvestment | 1        | .    | .    | .                       |
| _ce1 | lincome    | -.75     | .    | .    | .                       |
| _ce1 | lconsumption | -.0343804 | .0122816 | -2.80 | 0.005 | -.0584519 | -.010309 |
| _ce2 |          |           |       |      |                         |
| _ce2 | linvestment | -.1745742 | .00322 | -54.22 | 0.000 | -.1808852 | -.1682632 |
| _ce2 | lincome    | -.85     | .    | .    | .                       |
| _ce2 | lconsumption | 1        | .    | .    | .                       |

------------------------------------------------------------------------------
Summary

- Tip 1: Specifying the time structure
- Tip 2: Predictions with –arima-. Kalman Filter recursions
- Tip 3: Initial shock for Impulse response functions after -var-
- Tip 4: Unobserved component models with –sspace-
- Tip 5: Restrictions on cointegrating relationship for VEC models
Technical tips on time series with Stata

Gustavo Sánchez
Senior Statistician
StataCorp

2011 Mexican Stata Users Group Meeting 2011