Stochastic Integration

Gray Calhoun

November 4th, 2014, version 0.3.2

Plan for rest of semester

- Multivariate unit roots (three lectures)
  - Reading:
    - James Davidson's "Cointegration and Error Correction" (2012, *Handbook of Empirical Methods in Macroeconomics*)
    - Hamilton (1994) chapters 18 to 20
    - (Optional) Anna Mikusheva's lecture notes: 16 to 20
  - Lectures:
    1. Stochastic integration & spurious regression
    2. Cointegration, lecture 1
    3. Cointegration, lecture 2
- Dynamic Stochastic General Equilibrium models (three lectures)
  - Reading TBD
  - Lectures:
    1. State space models and the Kalman Filter
    2. DSGE models, lecture 1
    3. DSGE models, lecture 2
- Oh thank god, a break (equivalent to two lectures)
- Additional topics (two lectures on forecast evaluation)
- Scheduled exams

Quick review of persistent processes

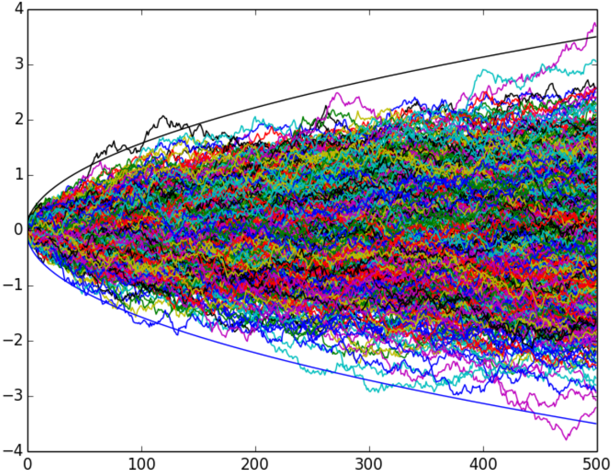
- Remember, if x_t is $I(0)$ then $\sum_{s=1}^{t} x_s$ is $I(1)$
  - We're going to focus on $I(1)$ processes, similar ideas hold for $I(2)$, etc.
- Simplest possible case: $x_t \sim MDS(0, \sigma^2)$, then

$$(1/\sqrt{T}) \sum_{s=1}^{[\lambda T]} x_t \Rightarrow \sigma W(\lambda)$$

- $W(\lambda)$ is a *Weiner process* or *Brownian Motion*, i.e.:
  1. it is continuous and mean zero
  2. $W(t) - W(s) \sim N(0, t - s)$ for any t and s
  3. Non-overlapping intervals are independent
- Quick graph of what this looks like:

```
using PyPlot
x = randn(500, 1000)
envelope = 3.5 * sqrt((1/500):(1/500):1)
plot([(cumsum(x,1)/sqrt(500)) envelope -envelope])
```

Simple graph of 1000 draws from Brownian Motion



Quick review of persistent processes

- We want to view each of the draws $W(\lambda)$ as a function of λ and view cumulative sums as approximate integrals

$$\sum_{t=1}^{[\lambda T]}(e_t/\sqrt{T}) \approx W(\lambda) = \int_0^{\lambda} dW(s)$$

- So $dW \approx e_t/\sqrt{T}$ and $\mathrm{E}\,dW^2 \approx \sigma^2/T$.
- Main building blocks: if $y_t = y_{t-1} + e_t$ with $e_t \sim MDS(0,\sigma^2)$ then
  - $(1/T)\sum_{t=2}^{T} y_{t-1}e_t \to^d \sigma^2 \int_0^1 W(t)dW(t) = (\sigma^2/2)(W(1)^2 - 1)$
  - $(1/T^{3/2})\sum_{t=2}^{T} y_{t-1} \to^d \sigma \int_0^1 W(t)dt$
  - $(1/T^2)\sum_{t=2}^{T} y_{t-1}^2 \to^d \int_0^1 W(t)^2 dt$
- You've seen with Helle that the OLS AR(1) coefficient for a univariate process is superconsistent and non-normal

$$T(\hat{\rho}-1) = \frac{(1/T)\sum_{t=2}^{T} y_{t-1}e_t}{(1/T^2)\sum_{t=2}^{T} y_{t-1}^2} \to^d \frac{W(1)^2 - 1}{2\int_0^1 W(t)^2 dt}$$

Quick review of persistent processes

- We can define more complicated stochastic integrals:

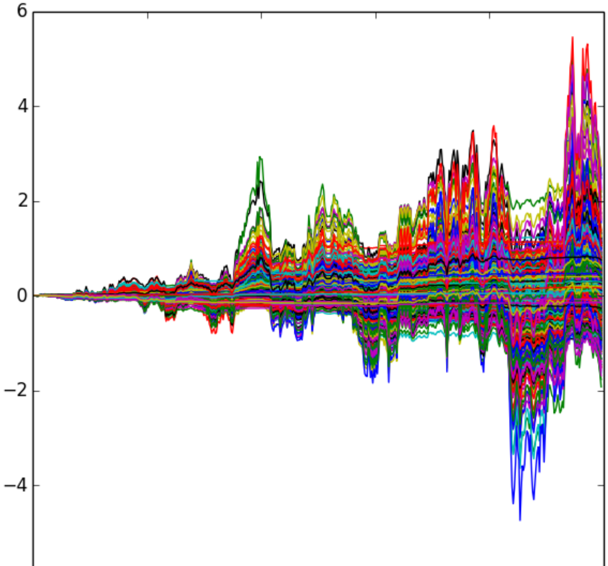
$$\int_0^\lambda W(s)ds + \int_0^\lambda W(s)^2 dW(s) =$$
$$\text{plim}(1/T)\sum_{t=2}^{[\lambda T]}\sum_{s=1}^{t-1}(e_s/\sqrt{T}) + \text{plim}\sum_{t=2}^{[\lambda T]}\Big(\sum_{s=1}^{t-1} e_s/\sqrt{T}\Big)^2 (e_t/\sqrt{T})$$

- Code for the graph

```
dW = randn(500, 1000) / sqrt(500)
W = cumsum(dW[1:499,:], 1)
plot(cumsum(W .* (1/500) .+ dW[2:500,] .* W.^2, 1))
```

- Note that $dW(s)$ is orthogonal to $W(s)$ (it takes place after)
- We only need a few basic results; there are entire classes you can take on working with Ito integrals and SDEs
- See White (2001) or Mikusheva (2013) for a little math and Davidson (1994) for much more math

Plot of our new stochastic integral



Additional points

- Everything continues to hold when $W(s)$ and $dW(s)$ are vectors
- Hamilton has a useful list of convergence results (Propositions 17.1 and 18.1)
- Continuous mapping theorem: if f is a continuous functional on $[0,1]$ then

$$f\Big(\sum_{t=1}^{[\lambda T]} e_t/\sqrt{T}\Big) \to^d f(W(\lambda))$$

  and, generally, if $g_t \to^d g$ where g is a random process on $[0,1]$ then $f(g_t) \to^d f(g)$
- Functional delta-method is similar
- Heteroskedasticity and autocorrelation affect the covariance process.
- Under (second order) stationarity, we replace σ^2 with the long-run variance of e_t. Without stationarity, it becomes more complicated.

Spurious regression

- Suppose that we have a bivariate unit root process

$$y_t = y_{t-1} + e_t = y_0 + \sum_{s=1}^{t} e_s$$

with $e_t \sim MDS(0, \sigma^2 I)$

- We run the regression

$$y_{1t} = \beta y_{2,t-1} + u_t$$

what happens?

- Note that we've set it up so that y_{1t} and $y_{2,t-1}$ should be unrelated.

Spurious regression

- Define
  - $v_1 = (1,0)'$
  - $v_2 = (0,1)'$
- The OLS coefficient can be written as

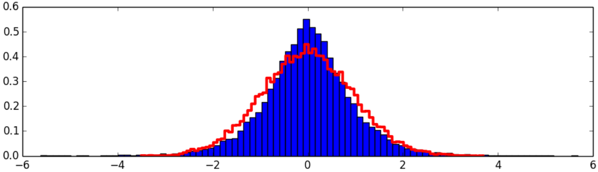
$$\hat{\beta} = \frac{\sum_{t=2}^{T} v_2' y_{t-1} \cdot y_t' v_1}{\sum_{t=2}^{T} v_2' y_{t-1} \cdot y_{t-1}' v_2} = \frac{v_2' \left(\frac{1}{T} \sum_{t=2}^{T} \frac{y_{t-1}}{\sqrt{T}} \frac{(y_{t-1}+e_t)'}{\sqrt{T}} \right) v_1}{v_2' \left(\frac{1}{T} \sum_{t=2}^{T} \frac{y_{t-1}}{\sqrt{T}} \frac{y_{t-1}'}{\sqrt{T}} \right) v_2}$$

$$= \frac{v_2' \left(\frac{1}{T} \sum_{t=2}^{T} \frac{y_{t-1}}{\sqrt{T}} \frac{y_{t-1}'}{\sqrt{T}} \right) v_1 + v_2' \left(\frac{1}{T} \sum_{t=2}^{T} \frac{y_{t-1}}{\sqrt{T}} \frac{e_t'}{\sqrt{T}} \right) v_1}{v_2' \left(\frac{1}{T} \sum_{t=2}^{T} \frac{y_{t-1}}{\sqrt{T}} \frac{y_{t-1}'}{\sqrt{T}} \right) v_2}$$

$$\Rightarrow \frac{v_2' \left(\int_0^1 W(s) W(s)' ds \right) v_1 + o_p(1)}{v_2' \left(\int_0^1 W(s) W(s)' ds \right) v_2}$$

- $\hat{\beta}$ is not consistent (i.e. doesn't converge to β)
- similar arguments show that $\sqrt{T}\hat{\beta}/\hat{\sigma}q$ diverges

Histogram of the distribution of $\hat{\beta}$




```
bh = Array(Float64, 20_000)
T = 500
@inbounds for i in 1:length(bh)
    W = cumsum(randn(T, 2), 1)
    bh[i] = sum(W[1:end-1,2].^2) \
                sum(W[1:end-1,2] .* W[2:end,1])
end
PyPlot.plt.hist(bh, 90, normed=1)

n = randn(20_000) * std(bh)
PyPlot.plt.hist(n, 90, normed = 1,
                histtype="step", linewidth=3)
```

Spurious regression

- Takeaway message: if you regress one unit root variable onto another, you will typically find significant nonzero coefficients whether or not there is any true relationship.
- Same intuition holds for regressing a unit-root process onto a trend.
- Same intuition holds for regressing a unit-root process onto a local trend.
- Some key papers
  - Granger and Newbold (1974) "Spurious regressions in econometrics"
  - Phillips (1986) "Understanding spurious regressions in econometrics"
  - Phillips and Durlauf (1986) "Multiple time series regressions with integrated processes"

Regression onto a stationary term

- Now suppose that we regress an I(1) process onto a covariance stationary I(0) regressor x_{t-1} (with mean zero)

$$y_t = \beta_0 x_{t-1} + \beta_1 + \beta_2 y_{t-1} + e_t$$

where β_2 is 1 but unknown.

  - assume that $\operatorname{var} e_t$ is 1 to keep the notation as simple as possible.
  - want to get limiting distributions for the OLS estimates

- A key problem: the different elements of $\hat{\beta} - \beta$ are going to converge at different rates.

$$\hat{\beta} - \beta = \Big(\sum_{t=2}^{T} (x_{t-1}, 1, y_{t-1})'(x_{t-1}, 1, y_{t-1}) \Big)^{-1} \sum_{t=2}^{T} (x_{t-1}, 1, y_{t-1})' e_t$$

we'll deal with this by rescaling the elements at different rates

Regression onto a stationary term

Let

$$\Lambda = diag(\sqrt{T}, \sqrt{T}, T)$$

so

$$\Lambda(\hat{\beta} - \beta) = \left(\Lambda^{-1} \sum_{t=2}^{T} \begin{pmatrix} x_{t-1}^2 & x_{t-1} & x_{t-1}y_{t-1} \\ x_{t-1} & 1 & y_{t-1} \\ x_{t-1}y_{t-1} & y_{t-1} & y_{t-1}^2 \end{pmatrix} \Lambda^{-1} \right)^{-1} \Lambda^{-1} \sum_{t=2}^{T} \begin{pmatrix} x_{t-1}e_t \\ e_t \\ y_{t-1}e_t \end{pmatrix}$$

$$\to^d \begin{pmatrix} \mathrm{E}x_t^2 & 0 & 0 \\ 0 & 1 & \int_0^1 W(s)ds \\ 0 & \int_0^1 W(s)ds & \int_0^1 W(s)^2 ds \end{pmatrix}^{-1} \begin{pmatrix} \sigma(\mathrm{E}x_t^2)^{1/2} W(1) \\ W(1) \\ (1/2)(W(1)^2 - 1) \end{pmatrix}$$

$$= \begin{pmatrix} \sigma(\mathrm{E}x_t^2)^{-1/2} W(1) \\ \begin{pmatrix} 1 & \int_0^1 W(s)ds \\ \int_0^1 W(s)ds & \int_0^1 W(s)^2 ds \end{pmatrix}^{-1} \begin{pmatrix} W(1) \\ (1/2)(W(1)^2 - 1) \end{pmatrix} \end{pmatrix}$$

- It's in nonstandard notation, but the estimator of $\hat{\beta}_0$ is normal with the usual variance.

Representative strategy for regressions with stationary and nonstationary terms

Suppose now you run the regression

$$y_t = \beta_0 + \beta_1 y_{t-1} + \beta_2 y_{t-2} + u_t$$

- we can rewrite the relationship as

  $$y_t = \beta_0 + \beta_1 \Delta y_{t-1} + (\beta_2 + \beta_1) y_{t-2} + u_t$$

  and estimating β_1 in this equation will give
  - A numerically identical estimate as in the levels equation
  - A consistent and asymptotically normal estimator of β_1
  - Note that our estimate of $\beta_1 + \beta_2$ will have an awkward distribution
- So the OLS estimate of β_1 in the original regression is consistent and asymptotically normal
- Similarly, we can show that the OLS estimate of β_2 in the original regression is consistent and asymptotically normal.
- Note that the estimate of β is not jointly normal, since $\beta_1 + \beta_2$ has a non-normal distribution.
- This is true whenever you can rewrite the expressions so that coefficients appear on I(0) components and has implications for cointegration.

Representative strategy for regressions with stationary and nonstationary terms

- In general, if you can rewrite the regression so that coefficients appear on stationary terms *simultaneously*, those coefficients will be jointly normal in the original regression.
- Key paper: Sims, Stock, and Watson (1990), "Inference in linear time series models with some unit roots"
- We'll see next time that cointegration complicates this

License and copying