OLS & MLE in EViews

Elmar Mertens

Study Center Gerzensee

AGENDA

- **EViews Basics**
- Linear Model and OLS
- OLS in EViews
- Maximum Likelihood
- Applications of ML in EViews

FILE & OBJECT MANAGEMENT IN EVIEWS

- Object oriented concept: `series`, `group`, `coef`, `equation`, `logl` ...
- All objects are stored in a "workfile"
- Most functions accessible via menus *and alternatively* via command line
- For descriptive stats on an object type "*myobject*`.stats`"
- Programs can be scripts or functions. They are kept in separate files

CREATING A SERIES IN EVIEWS

BACKUP

`series mynewseries = expression`, for example:

- Returns on a series of equity prices "`price`"
  - `series simplereturns = ...`
    `(price - price(-1)) / price(-1)`
  - `series logreturns = ...`
    `log(price) - log(price(-1))`
- Generate random numbers:
  `series y =` *mymean* `+` *myvol* `* nrnd`
- Dummies: `series mydummy = y > 0`

DATA IMPORT

BACKUP

- Can import ASCII, Excel and more
- First, need to open/create an appropriately sized workfile ("File" $\rightarrow$ "New" $\rightarrow$ "Workfile")
- Second, open import dialogue and make your choices ("File" $\rightarrow$ "Import" $\rightarrow$ ...)
- Can do same with commands *`workfile`*, *`new`* , *`create`* and *`read`*. (Care about exports? Try *`write`*.)
- Example program: `ffread.prg`

EXERCISE: TIME SERIES

`MAKERETURNS.PRG, GENRAR1.PRG`

i) Open workfile `sp500prices.wf1` with daily closing prices of the S&P 500

ii) Compute simple $R_t = \frac{P_t - P_{t-1}}{P_{t-1}}$ and log-returns $r_t = \log P_t / P_{t-1}$ on the S&P 500

iii) Generate data on $x \sim iidN(3, 20)$

iv) Generate data on a time series y following an AR(1), such that it has mean $\mu = 2$ and autocorrelation $\rho = 0.8$.

Recall: $y_t = (1-\rho)\mu + \rho y_{t-1} + \varepsilon_t$, $\varepsilon_t \sim iidN(0, 1)$

EViews Hint: First set an initial value for y, then operate on the rest of the sample "`smpl @first + 1 @last`"

EXERCISE SOLUTIONS

- You find solutions to each exercise in the programs indicated at the top of each exercise slide
- For the previous exercise these were `makereturns.prg` and `genrar1.prg`
- Typically, these are small scripts, whose commands could also be entered line by line on the command line
- Open them with "File" $\rightarrow$ "Open" $\rightarrow$ "Program"
- Important: Select "Update Default Directory" in the dialog box
- Hit the "Run" button in the program's window

SAMPLE MANAGEMENT

- Either "on the fly" *smpl* 1980:1 1999:12

  Or with pre-defined sample objects:

  ```
  sample mysample 1980:1 1999:12
  smpl mysample
  ```

- Neat are the placeholders `@first`, `@last`, `@all`:

  ```
  smpl @first + 1 @last
  smpl @all
  ```

- Conditional Samples:

  ```
  smpl if returns > 0
  ```

- Note: EViews adjusts itself automatically to missing values / initial values (e.g. for AR-estimation)

EXERCISE: CONDITIONAL SAMPLES

RECESSIONRETURNS.PRG

i) Open workfile `ff.wf1` with monthly U.S. stock returns from Fama/French.

ii) Look at the Market Return in Excess of Riskfree `rmrf` and the indicator variable `recession`. The latter is equal to one if there was a NBER recession.

iii) What is the average return and standard deviation over the full sample?

iv) How do these statistics look if you condition on being in a recession or a boom?

AGENDA

- EViews Basics
- **Linear Model and OLS**
- OLS in EViews
- Maximum Likelihood
- Applications of ML in EViews

LINEAR REGRESSION MODEL

Key Assumptions:

- Linearity: $y_i = \boldsymbol{x}_i' \boldsymbol{\beta} + \varepsilon_i$
- Orthogonality: $\mathrm{E}\left[\boldsymbol{x}_i \varepsilon_i\right] = \mathbf{0}$, a.k.a. "predetermined"
- To get standard errors, some Central Limit Theorem (CLT) should also hold for $\boldsymbol{g}_i = \boldsymbol{x}_i \varepsilon_i$:

$$\sqrt{n}\left(\frac{1}{n}\sum_{i=1}^{n} \boldsymbol{g}_i\right) \rightarrow N(\mathbf{0}, \boldsymbol{S})$$

  If $\boldsymbol{g}_i$ serially independent, $\boldsymbol{S} = \mathrm{E}\left[\boldsymbol{g}_i \boldsymbol{g}_i'\right]$. Otherwise autocovariances of $\boldsymbol{g}_i$ matter, too (a.k.a. "spectral density at frequency zero", estimate with "HACCME")

JUSTIFICATIONS FOR OLS ESTIMATOR

1. **Least Squares**:

$$\boldsymbol{b} = \underset{\boldsymbol{\beta}}{\operatorname{argmin}} \sum_{i=1}^{n} \left(y_i - \boldsymbol{x}_i' \boldsymbol{\beta}\right)^2$$

2. **Method of Moments (GMM)**: Sample analogue to

$$\mathrm{E}\left[\boldsymbol{x}_i \left(y_i - \boldsymbol{x}_i' \boldsymbol{\beta}\right)\right] = 0$$

3. **Maximum Likelihood**: assuming $\varepsilon_i \sim iidN(0, \sigma^2)$

- All deliver the same estimate! (but s.e. may differ)

THE ESTIMATOR

BACKUP

True coefficients

$$\boldsymbol{\beta} = \mathrm{E}\left[\boldsymbol{x}_i \boldsymbol{x}_i'\right]^{-1} \mathrm{E}\left[\boldsymbol{x}_i y_i\right]$$

estimated by sample analogues (LLN):

$$\boldsymbol{b} = \left(\frac{1}{n}\sum_{i=1}^{n} \boldsymbol{x}_i \boldsymbol{x}_i'\right)^{-1} \frac{1}{n}\sum_{i=1}^{n} \boldsymbol{x}_i y_i$$

CLT yields asymptotic distribution of estimator:

$$\sqrt{n}\,(\boldsymbol{b} - \boldsymbol{\beta}) \to N\left(\mathbf{0}, \mathrm{E}\left[\boldsymbol{x}_i \boldsymbol{x}_i'\right]^{-1} \boldsymbol{S}\,\mathrm{E}\left[\boldsymbol{x}_i \boldsymbol{x}_i'\right]^{-1}\right)$$

If homoscedastic: $\boldsymbol{S} = \mathrm{var}(\varepsilon_i) \cdot \mathrm{E}\left[\boldsymbol{x}_i \boldsymbol{x}_i'\right]$, otherwise HACCME

LLN & CLT

BACKUP

Place some (weak) restrictions on dependence, heterogeneity and existence of moments for stationary data $\boldsymbol{g}_i$:

Law of Large Numbers: "Consistency of sample mean"

$$\frac{1}{n}\sum_{i=1}^{n}\boldsymbol{g}_i \to \mathrm{E}\left[\boldsymbol{g}_i\right]$$

Central Limit Theorem: "*Asymptotic* distribution"

$$\sqrt{n}\left(\frac{1}{n}\sum_{i=1}^{n}\boldsymbol{g}_i - \mathrm{E}\left[\boldsymbol{g}_i\right]\right) \to N(\mathbf{0}, \boldsymbol{S})$$

Linear independence: $\boldsymbol{S} = \mathrm{E}\left[\boldsymbol{g}_i\boldsymbol{g}_i'\right]$

Else: Spectral density at frequency zero

STATIONARITY

- Key assumption is stationarity, in particular

$$Ex^2 = \text{const} < \infty$$

- Non-stationary example: Random Walk:

$$\begin{aligned} s_t &= s_{t-1} + r_t \\ r_t &= \bar{r} + \sigma\varepsilon_t \qquad \varepsilon_t \sim \ iid(0,1) \\ \text{Var}_t s_{t+1} &= \sigma^2 \\ \text{Var}_t s_{t+k} &= k\sigma^2 \to \infty \end{aligned}$$

- Ever wondered why "Asset *Pricing*" is so obsessed with *returns* ;-)

STATIONARITY TESTS

- Basic idea: $y_t = \rho y_{t-1} + BX_{t-1} + e_t \qquad \rho = 1$??
- Null hypothesis of $\rho = 1$ says "y_t has unit root"
- Can run OLS but need special s.e. for tests of ρ (but not for B), e.g. "Dickey-Fuller", "Ng-Perron"
- Problem: In finite sample we can never really distinguish between $\rho = 1$ and $\rho = 0.99\ldots$. Low Power of tests. Check with null of "y_t is stationary" (KKSS test)
- Can interest rates be unit root? Inflation? Hours worked?
- Always ask "What is better approximation for this sample"

AGENDA

- EViews Basics
- Linear Model and OLS
- **OLS in EViews**
- Maximum Likelihood
- Applications of ML in EViews

OLS IN EVIEWS

- Quick: Just list dependent and explanatory variables:

```
ls y c x1 x2
```

- Or stored in a named equation object:

```
equation myequation.ls y c x1 x2
myequation.stats
```

Some Tips:

- Group regressors: "`ls y c xgroup`" (neat for programs)
- Freeze results in separate table: "`freeze myequation.stats`"
- Non-linear LS: "`ls y = c(1) + x1^c(2)` + c(3)*x2" estimates c_0,c_1 and c_2 in $y_t = c_0 + x_{1,t}^{c_1} + c_2 x_{2,t} + \varepsilon_t$

ESTIMATION OPTIONS

BACKUP

- Variance/Covariance Matrices:
  - "Huber-Eicker-White", a.k.a. "HCCME": `ls(h)`
  - "Newey-West", a.k.a. "HACCME": `ls(n)`
- Other Options:
  - Weighted Least Squares with given weights
    `ls(w=myweights)`
  - Optimization methods, might matter for non-linear least squares. For details see manual:
    "Help" → "Quick Help Reference" → "Object Reference" → "Equation" → "ls"

EXERCISE: REGRESSIONS

`FFXR.PRG, FFMM.PRG`

i) Open workfile `ff.wf1` with returns from Fama & French

ii) Look at descriptive statistics of market's return in excess of the riskfree rate, R_t^M, stored in `rmrf`

iii) Create returns in excess of `rf` for `r1`: $R_t^{e,1} = R_t^1 - R_t^f$

iv) Run market model regression of `r1`'s excess return on constant and market: $R_t^{e,1} = \alpha + \beta R_t^M + \varepsilon_t$

v) Run "three factor model" regression of `r1`'s excess return on constant, market (`rmrf`), "size" (`smb`) and "value" (`hml`): $R_t^{e,1} = \alpha + \beta_1 R_t^M + \beta_2 R_t^{SMB} + \beta_3 R_t^{HML} + \varepsilon_t$

EXERCISE: UNIT ROOTS

`INTROOTS.PRG`

i) `interestrates.wf1` contains monthly TBill and Funds Rate for the U.S.

ii) Test for stationarity using Dickey-Fuller and other tests
Hint: Open series and use “View” $\rightarrow$ “Unit Root Tests” or `tb3.uroot(dfgs)`

iii) Test over different samples

EXERCISE: WALD & ARCH

`SPGARCH.PRG`

i) Go back to `sp500prices` and estimate an AR(7) for the log-returns

ii) Are any of the individual coefficients significant?

iii) Test whether all of the AR(7) coefficients are jointly significant

iv) Test for presence of heteroscedasticity (Hint: use `archtest`)

v) If you find heteroscedasticity, were the estimated AR(7) coefficients actually valid?

vi) Estimate an ARCH/GARCH model using the AR(7) as mean equation

EQUATION SYSTEMS

BACKUP

- Create *system* object: *`system`* `mysystem`
- Open its "specification window": `mysystem.`*`spec`*
- Simply list the set of equations in "longhand" notation:

```
xr1 = c(1) + c(2) * rmrf
' or nicer using named coeffs:
xr1 = alpha(1) + beta(1) * rmrf
```

- Assuming independence between equations' *residuals*: `mysystem.`*`ls`* otherwise: `mysystem.`*`sur`*

EXERCISE: REGRESSION SYSTEMS

`FFGRS.PRG, BJ.PRG`

i) Run a system of market model regressions using all six portfolios `xr1 ...xr6`:

$$R_t^{e,1} = \alpha_1 + \beta_1 R_t^M + \varepsilon_t^1$$

$$\vdots$$

$$R_t^{e,6} = \alpha_6 + \beta_6 R_t^M + \varepsilon_t^6$$

ii) Run a regression with 1 as dependent and each of `rmrf, r1, r2 ... r6`'s excess return as explanatory variable without constant.

$$1 = \beta_1 R_t^{e,1} + \beta_2 R_t^{e,2} + \ldots + \beta_6 R_t^{e,6} + \beta_M R_t^M + \varepsilon_t$$

TESTING LINEAR RESTRICTIONS

- Britten-Jones (1999) shows that the following regression yields weights of risky assets in mean-variance tangency portfolio

$$1 = \beta_1 R_t^{e,1} + \beta_2 R_t^{e,2} + \ldots + \beta_6 R_t^{e,6} + \beta_M R_t^M + \varepsilon_t$$

- If `rmrf` is tangency, all other coefficients should be zero
- But individual t-stats can be misleading (Gibbons, Ross, and Shanken 1989). Better do a joint "Wald" test instead

BJ.PRG

```
series ones = 1
equation bj.ls ones xr rmrf
bj.wald c(1) = 0, c(2)=0, c(3)=0, c(4)=0, c(5)=0, c(6)=0
```

Note: Wald tests work the same with *system* or *logl* object

EXERCISE: WALD TESTS

`FFGRS.PRG, BJ.PRG`

i) Test the restrictions of Britten-Jones (1999) using the six excess returns (`xr1`, ... `xr6`) and the market `rmrf`

ii) For the system estimated in the previous exercise, test the restriction that all intercepts are *jointly* zero:

$$\alpha_1 = 0, \alpha_2 = 0, \alpha_3 = 0, \alpha_4 = 0, \alpha_5 = 0, \alpha_6 = 0$$

iii) Compare the Test Statistics, what do you observe?

AGENDA

- EViews Basics
- Linear Model and OLS
- OLS in EViews
- **Maximum Likelihood**
- Applications of ML in EViews

MOTIVATION FOR MAXIMUM LIKELIHOOD

- Generic approach with well-known and mostly desirable properties
  - Consistent and efficient estimator
  - Invariance Principle: $\boldsymbol{\theta}$ vs $f(\boldsymbol{\theta})$
  - Numerical procedures (now) cheap and ubiquitous
- Straightforward concept *if* distribution of data is known
- Can be understood as particular GMM, this perspective allows for some robustness against misspecification

GENERAL FRAMEWORK FOR ML

- Suppose that the iid data $\boldsymbol{w}_i$ was generated from a distribution whose pdf is parametrized by $\boldsymbol{\theta}$, $f(\boldsymbol{w}_i;\boldsymbol{\theta})$

- For which $\boldsymbol{\theta}$ were *fixed* $\boldsymbol{w}_i$ most likely observed?

$$\boldsymbol{\theta}^{\mathsf{MLE}} = \underset{\boldsymbol{\theta}}{\operatorname{argmax}} \sum_{i=1}^{n} \log f(\boldsymbol{w}_i;\boldsymbol{\theta})$$

$$\Rightarrow \qquad \sum_{i=1}^{n} \frac{\partial \log f(\boldsymbol{w}_i;\boldsymbol{\theta}^{\mathsf{MLE}})}{\partial \boldsymbol{\theta}} = 0$$

Jargon:	$\log f(\boldsymbol{w}_i;\boldsymbol{\theta})$	a.k.a.	"contributions"
	$\frac{\partial \log f(\boldsymbol{w}_i;\boldsymbol{\theta}^{\mathsf{MLE}})}{\partial \boldsymbol{\theta}}$	a.k.a.	"scores"

LINEAR REGRESSION WITH NORMAL ERRORS

Assuming $\varepsilon_i \sim iidN(0, \sigma^2)$ in $y_i = \boldsymbol{x}_i'\boldsymbol{\beta} + \varepsilon_i$:

$$f(\underbrace{y_i, \boldsymbol{x}_i}_{\boldsymbol{w}_i}; \underbrace{\boldsymbol{\beta}, \sigma}_{\boldsymbol{\theta}}) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(y_i - \boldsymbol{x}_i'\boldsymbol{\beta})^2}{2\sigma^2}\right]$$

Hence the log-likelihood objective is

$$\frac{1}{n}\sum_{i=1}^{n} \log f(y_i, \boldsymbol{x}_i; \boldsymbol{\beta}, \sigma) =$$

$$-\frac{1}{2}\log(2\pi) - \frac{1}{2}\log\sigma^2 - \frac{1}{2}\underbrace{\frac{1}{n}\sum_{i=1}^{n}\left(\frac{y_i - \boldsymbol{x}_i'\boldsymbol{\beta}}{\sigma}\right)^2}_{\text{Least Squares !}}$$

STATISTICS OF NORMAL DISTRIBUTION

BACKUP

Assuming $y_i \sim iidN(\mu, \sigma^2)$:

$$f(\underbrace{y_i}_{\boldsymbol{w}_i}; \underbrace{\mu, \sigma}_{\boldsymbol{\theta}}) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(y_i - \mu)^2}{2\sigma^2}\right]$$

Hence the log-likelihood objective is

$$\frac{1}{n}\sum_{i=1}^{n} \log f(y_i; \mu, \sigma) =$$

$$-\frac{1}{2}\log(2\pi) - \frac{1}{2}\log\sigma^2 - \frac{1}{2}\frac{1}{n}\sum_{i=1}^{n}\left(\frac{y_i - \mu}{\sigma}\right)^2$$

"ML IS GMM ON THE SCORES" BACKUP

- In analogy to GMM with $\boldsymbol{g}_i = \frac{\partial \log f(\boldsymbol{w}_i;\boldsymbol{\theta}^{\mathsf{MLE}})}{\partial \boldsymbol{\theta}}$ we get

$$\sqrt{n}\left(\boldsymbol{\theta}^{\mathsf{MLE}} - \boldsymbol{\theta}_0\right) \rightarrow N\left(\mathbf{0}, \boldsymbol{A}^{-1}\boldsymbol{S}\boldsymbol{A}^{-1}\right)$$

$$\boldsymbol{S} = \mathrm{E}\left[\boldsymbol{g}_i \boldsymbol{g}_i'\right] = \mathrm{E}\left[\frac{\partial \log f}{\partial \boldsymbol{\theta}} \frac{\partial \log f}{\partial \boldsymbol{\theta}}'\right]$$

$$\boldsymbol{A} = \mathrm{E}\left[\frac{\partial \boldsymbol{g}_i}{\partial \boldsymbol{\theta}'}\right] = \mathrm{E}\left[\frac{\partial^2 \log f}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'}\right] \text{ (}\boldsymbol{S}\text{ and }\boldsymbol{A}\text{ evaluated at }\boldsymbol{\theta}_0\text{)}$$

- ML's Information Matrix Equality: $\boldsymbol{S} = -\boldsymbol{A}$
- Cramer-Rao: No tighter s.e. possible than $\boldsymbol{S}$! (or bias)
- If not sure whether pdf really true but $\boldsymbol{g}_i$ still a sensible moment: Quasi-ML uses robust estimation of $\boldsymbol{A}$ and $\boldsymbol{S}$

AGENDA

- EViews Basics
- Linear Model and OLS
- OLS in EViews
- Maximum Likelihood
- **Applications of ML in EViews**

ML IN EVIEWS

Name of the Game: Given $\boldsymbol{\theta}$, describe contributions as a series

1. Create log-likelihood object: `logl mylogl`

2. In the object's specification window (`mylogl.spec`):

```
@logl mycont
mycont = -log(c(2)) - (mydata - c(1))^2/c(2)
```

$\underbrace{\texttt{-log(c(2)) - (mydata - c(1))\^{}2/c(2)}}_{\text{Example of normal's mean and variance}}$

3. Estimate: `mylogl.ml`

If you run into errors or strange results, use some sensible starting values for your coefficients!

COIN FLIPPING

- When flipping a coin we have

$$\begin{aligned} \text{Prob(Heads)} &= \delta \\ \text{Prob(Tails)} &= 1 - \delta \end{aligned}$$

- The coin might be loaded hence $\delta \epsilon [0; 1]$

- Collect data on draws: $w_i = \begin{cases} 1 & \text{if Heads} \\ 0 & \text{if Tails} \end{cases}$

Note: The distribution of a single draw is a.k.a. "Bernoulli". If you count the number of heads in a sequence of draws, the sum is distributed "Binomial"

MLE OF COIN PROBABILITY IN EVIEWS

- Individual flip has *pdf*

$$f(w_i;\delta) = \delta^{w_i}(1-\delta)^{1-w_i}$$

- Recall: MLE of δ maximizes *log*-likelihood of the n independent draws

$$\delta^{\mathsf{MLE}} = \underset{\delta}{\operatorname{argmax}} \sum_{i=1}^{n} \underbrace{\log f(w_i;\delta)}_{\text{“contribution”}}$$

- In Eviews:

COINFLIPS.WF1

```
@logl llf
llf = log(c(1)^w * (1-c(1))^(1-w))
```

COIN PROBABILITY WITH OLS/GMM

BACKUP

- MLE yields sample average. This is no coincidence. Notice:

$$\mathrm{E}[w_i] = 1 \cdot \delta + 0 \cdot (1 - \delta) = \delta$$

- Method of moments:

$$\hat{\delta} = \frac{1}{n} \sum_{i=1}^{n} w_i$$

- Or think of a linear regression $w_i = \delta + \varepsilon_i$ The errors are non-normal, but they satisfy the orthogonality condition $\mathrm{E}[\varepsilon_i] = 0$. LLN and CLT can be applied.

- In EViews: "`ls w c`" or "`w.stats`" again see `coinflips.wf1`

EXERCISE: TAILS OF DISTRIBUTION

`TAILSMLE.PRG`

Use `sp500prices.wf1`:

i) Create log-returns

ii) Estimate the probability of a return being

- lower than -1%
- higher than 1%
- being between -1% and 1%

Hint: For each of the conditions above, create a series which takes either zero or one, depending on whether condition is true. For example `series low = logreturns < -.01`

iii) Do you have to use MLE for this?

EXTENSIONS TO LOGL-OBJECT

BACKUP

Initial Values In `logl.spec`: "`@param c(1) .5 c(2) 1`" to start optimization with `c(1)= 0.5` and `c(2)=1`

Named Coefficients Instead of `c(1), c(2)` etc.: Define your own coefficient vectors on command line for use in logl or equation objects: `coef(`*n*`) mycoef`

Intermediate series For clarity of logl specification, e.g.

```
e = (data - mu(1)) / sigma(1)
logl = -e^2 - log(sigma(1)^2)
```

Analytic Scores Sometimes easy to derive or given in textbooks. Very useful for speed and accuracy. (See EViews Manual)
Quiz: What's the analytic score for a Linear Regression?

EXERCISE: SIMPLE MLE

`MEANVAR.PRG, LINREG.PRG`

Still using `ff.wf1`:

i) Estimate mean and variance of portfolio 1's excess return `xr1` assuming normality

ii) Redo market model of `xr1` on constant and `rmrf` using maximum-likelihood. Compare results to previous OLS estimation

Hint: To use the same `logl` for differing dependent and explanatory variables, express everything in terms of series `y` and `x` and assign data prior to estimation

SOME FUNCTIONS IN EVIEWS

BACKUP

- `log(·)` natural logarithm
- `@sqrt(·)` square root
- `@dnorm(·)` density of standard normal
- `series myseries = @recode(`*`cond, expr1, expr2`*`)` Fills series with *expr1* if *cond* true, otherwise with *expr2* (Useful for AR1 later)

See also "Help" → "Quick Help Reference" → "Function Reference"

AR(1) ESTIMATION

With $\varepsilon \sim iidN\left(0, \sigma^2\right)$ the AR(1) model

$$y_t - \mu = \rho(y_{t-1} - \mu) + \varepsilon_t \quad (|\rho| < 1)$$

implies unconditional *and* conditional distributions for y:

$$y_t \sim N\left(\mu, \frac{\sigma^2}{1-\rho^2}\right) \quad y_{t|t-1} \sim N\left(\mu(1-\rho) + \rho y_{t-1}, \sigma^2\right)$$

OLS takes y_1 as given and is a.k.a. *conditional* MLE. Hence, its log-likelihood consists of $T-1$ contributions. Exact ML adds a contribution for y_1 based on its unconditional distribution.

Quiz: What happened to our original ML-assumption of iid data?

LOG-LIKELIHOOD OF NORMAL-AR(1)

BACKUP

$$\sum_{t=1}^{T} \log f_t(\boldsymbol{y};\boldsymbol{\theta}) = \underbrace{\log \phi\left(\frac{y_1-\mu}{\sqrt{\sigma^2/(1-\rho^2)}}\right) - \frac{1}{2}\log\left(\frac{\sigma^2}{1-\rho^2}\right)}_{\text{contribution of } y_1}$$

$$+ \underbrace{\sum_{t=2}^{T}\left\{\log\phi\left(\frac{(y_t-\mu)-\rho(y_{t-1}-\mu)}{\sigma}\right) - \frac{1}{2}\log(\sigma^2)\right\}}_{\text{contributions of } y_2 \ldots y_T\text{, constitute OLS/conditional MLE}}$$

Where $\boldsymbol{y} = (y_1 \ldots y_T)'$, $\boldsymbol{\theta} = (\mu, \rho, \sigma)'$ and $\phi(x) = \frac{\exp(-x^2/2)}{\sqrt{2\pi}}$ is the standard normal's *pdf*

EXERCISE: AR(1) WITH MLE

`GENRAR1.PRG,AR1.PRG`

i) Simulate a series `y` as AR(1) process

ii) Estimate *conditional* AR(1) with OLS
(You might also want to check whether your MLE code of the linear model still works here. Why shouldn't it?)

iii) Now create the exact MLE for this AR(1) and re-estimate the parameters

Hint: When creating a series, you can specify lags of a series `y` simply as `y(-1)`

FURTHER APPLICATIONS OF ML

ARCH/GARCH models with various distributions. For non-normal or threshold models in EViews: `logl`

Kalman Filters when the filtering coefficents are unknown, for example see Hamilton (1994, Chapter 13)

Choice Models for example binary ("yes/no") or multinomial ("coffee, tea, ardent spirits"). Binary logit/probit is canned in EViews. For multinomial, need `logl` again. Cool: The logit can equivalently be set up as GMM in EViews (Greene 2000, p. 821/860)!

References

Britten-Jones, Mark. 1999. "The Sampling Error in Estimates of Mean Variance Efficient Portfolio Weights." *The Journal of Finance* 54 (2): 655–671 (April).

Fama, Eugene F., and Kenneth R. French. 1993. "Common Risk Factors in the Returns on Stocks and Bonds." *The Journal of Financial Economics* 33:3–56.

Gibbons, Michael R., Stephen A. Ross, and Jay Shanken. 1989. "A Test of the Efficiency of a Given Portfolio." *Econometrica* 57 (5): 1121–1152 (September).

Greene, William H. 2000. *Econometric Analysis*. 4th edition. Upper Saddle River, NJ: Prentice Hall.

Hamilton, James D. 1994. *Time-Series Analysis*. Princeton, NJ: Princeton University Press.

Hayashi, Fumio. 2000. *Econometrics*. Princeton, NJ: Princeton University Press.

Newey, W.K., and K.D. West. 1987. "A simple positive semi-definite heteroskedasticity and autocorrelation consistent covariance matrix." *Econometrica* 55:703–708.

White, Halbert. 1980. "A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity." *Econometrica* 48:817–837.

———. 2001. *Asymptotic Theory for Econometricians*. Revised Edition. New York: Academic Press.