State Space Models

Advanced Time Series Econometrics

Scottish Graduate Programme in Economics

Introduction

- You probably have had some earlier study of of time series econometrics
- Two main concepts there “regression and trends”
- Distributed lag models are regressions
- Autoregressive models = regression where explanatory variables dependent variables
- VAR = regressions with many equations
- Even ARCH and GARCH equations for error variances are similar to regressions
- Trends = unit roots, cointegration/error correction involves variables trending together, etc.
- Thinking in terms of regressions + trends can get you pretty far
- From your earlier study you have a good set of tools to do many things

Introduction

- State space modelling is a different way of approaching time series econometrics
- Think of time series variables being driven by states
- State space models can do things similar to regressions and trends, plus much more
- A few themes:
- Trends as states
- Parameters changing (structural breaks, regime switching, time-variation) as states
- Factor methods: Information in Big Data (100+ variables) summarized in terms of a few states
- Missing values and mixed frequencies

Readings

- Ghysels and Marcellino chapter 11 offers a good intro to general state space models
- Withiin this general class, today I will lecture on structural time series (STS) models
- Tsay chapter 11 has more on STS
- Classic state space books (which cover much more than we do in this course):
- Durbin, J. and Koopman, S. (2012)., Time Series Analysis by State Space Methods (second edition).
- Harvey, A. (1993) Time Series Models.
- Harvey, A. (1989) Forecasting, Structural Time Series Models and the Kalman Filter.
- Prado, R. and West, M. (2010). Time Series: Modelling, Computation and Inference.
- West, M. and Harrison, J. (1997). Bayesian Forecasting and Dynamic Models (second edition).

What Can State Space Models Do?

- Models mentioned above: STS, factor models, models with parameter change, mixed frequency models
- But also many other things:
- Stochastic volatility (an alternative to GARCH)
- Dynamic stochastic general equilibrium (DSGE) models are state space models
- Splines (not covered in this course)
- Continuous time models (not covered in this course)

Econometrics of State Space Models

- An introductory econometrics course will cover standard set of statistical tools for estimation, hypothesis testing and forecasting
- E.g. least squares methods, maximum likelihood
- Programs like R allow for estimation in practice
- State space models same idea (usually maximum likelihood)
- Very little discussion of estimation will be given (in practice, use R)
- I want to focus on the models, their properties and using them in practice

The Local Level (Local Trend) Model

- Explain basic ideas in simplest state space model: the local level model
- For an observed variable, y_t, $t = 1, .., T$ have

$$y_t = \alpha_t + \varepsilon_t$$

- ε_t is i.i.d. $N\left(0, \sigma_\varepsilon^2\right)$.
- α_t which is not observed (called a *state*) and follows random walk for $t = 1, .., T - 1$:

$$\alpha_{t+1} = \alpha_t + u_t$$

- u_t is i.i.d. $N\left(0, \sigma_u^2\right)$
- ε_t and u_s are independent of one another for all s and t.
- First equation: measurement (observation) equation, second state equation
- α_1 is *initial condition*.

Relationship to Other Models

- Can write local level model as

$$\Delta y_t = \varepsilon_t - \varepsilon_{t-1} + u_{t-1}$$

- Note moving average (MA) structure for errors
- Δy_t is stationary ($I(0)$) whereas y_t has unit root ($I(1)$)
- Can write

$$\alpha_t = \alpha_1 + \sum_{j=1}^{t-1} u_j$$

- this is a trend (stochastic trend)
- local level model decomposes y_t, into a trend component, α_t, and an error or irregular component, ε_t.
- Test of whether $\sigma_u^2 = 0$ is one way of testing for a unit root.
- Note: if $\sigma_u^2 = 0$ then $u_t = 0$ and $\alpha_{t+1} = \alpha_t$ for all t (trend vanishes, just becomes a constant intercept)
- All usual univariate time series things: ARIMA modelling, unit root testing, etc. can be done in state space framework

Relationship to Other Models

- α_t is the mean (or level) of y_t.
- Mean is varying over time, hence terminology *local level model*
- Measurement equation can be interpreted as simple example of regression model involving only an intercept.
- But the intercept varies over time: time varying parameter model
- Easy to add explanatory variables to create regression model with time varying intercept
- Extensions of local level model used to investigate parameter change in various contexts.

Example: Trend Inflation

- Central bankers often interested in measures of trend (underlying) inflation
- Decompose observed inflation (π_t) into permanent (π_t^*) and transitory (c_t) components:

$$\pi_t = \pi_t^* + c_t.$$

- π_t^* is underlying inflation, defined through the properties:

$$\begin{aligned} E_t\left(\pi_{t+h}\right) &\rightarrow E_t\left(\pi_{t+h}^*\right) \\ E_t\left(c_{t+h}\right) &\rightarrow 0 \text{ as } h \rightarrow \infty. \end{aligned}$$

- Do not observe π_t^*, but effectiveness of monetary policy can depend crucially on it
- Central bankers distinguish between anchored, contained and unmoored inflation expectations
- Different econometric models developed based on these concepts

Anchored inflation expectations

- Assume a credible inflation target exists, $\overline{\pi}$
- Many researchers model underlying inflation as:

$$\pi_t^* = \overline{\pi}\,(1 - \theta) + \theta \pi_{t-1}^* + u_t$$

- u_t is a stationary residual and $|\theta| < 1$.
- Can be shown that trend inflation will be pulled back to the inflation target
- This model has anchored inflation expectations
- Inflation expectations perfectly anchored if $\theta = 0$

Unmoored inflation expectations

- Many (e.g. Stock and Watson, 2007, JMCB) model underlying inflation as a random walk:

$$\pi_t^* = \pi_{t-1}^* + u_t.$$

- A shock (u_t) hits trend inflation it has permanent effect
- This is property of random walk
- $$\pi_t^* = \sum_{j=1}^{t} u_j$$
- Each past shock (u_j for $j = 1, .., t$) appears in formula (permanent effect)
- Unmoored inflation expectations

How Does This Relate to State Space Models?

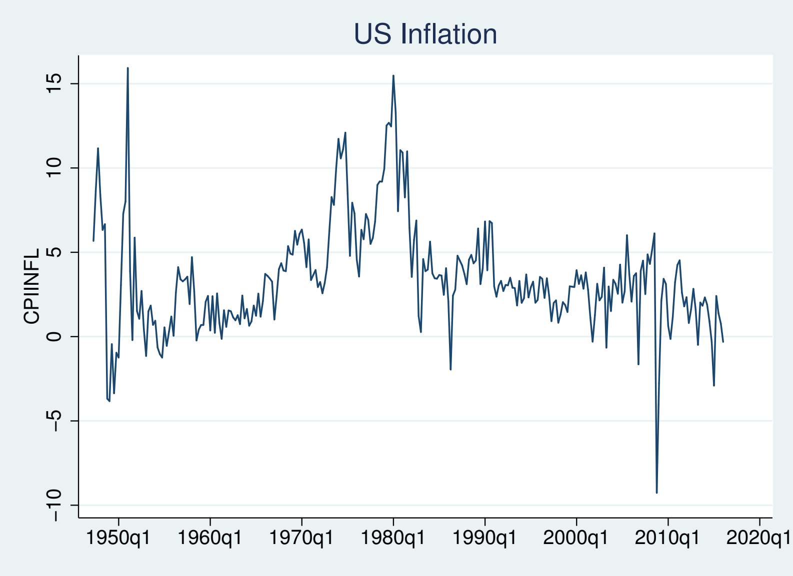
- Trend inflation, π_t^*, is the state (α_t)
- Inflation, y_t, is observed
- Unmoored model of Stock and Watson is the local level model
- Anchored inflation expectations model is a slight generalization
- One variant might be to have state equation:

$$\alpha_{t+1} = c + \rho\alpha_t + u_t$$

  where c is constant and ρ a coefficient to be estimated
- Key point: All state space models

Example: US Inflation

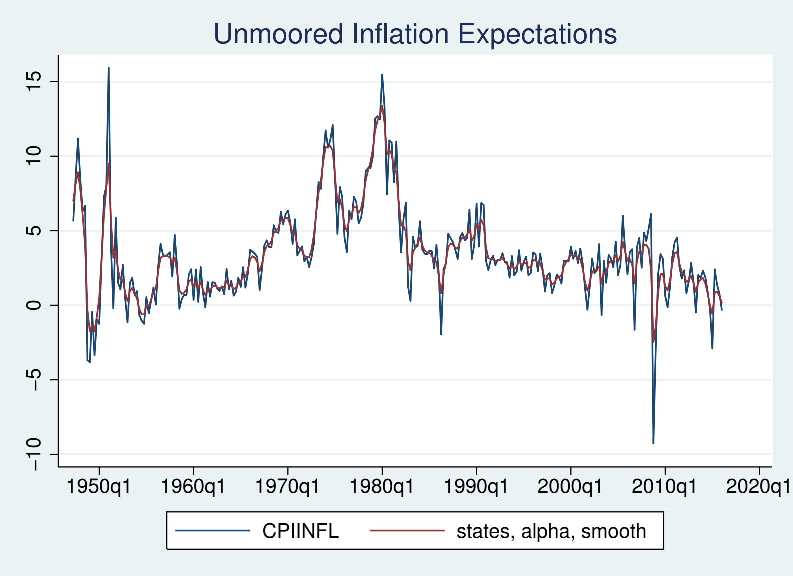
- Next figure graphs US CPI inflation
- Quarterly changes made into annual rate
- If P_t is CPI in quarter t
- $400 \ln\left(\frac{P_t}{P_{t-1}}\right)$



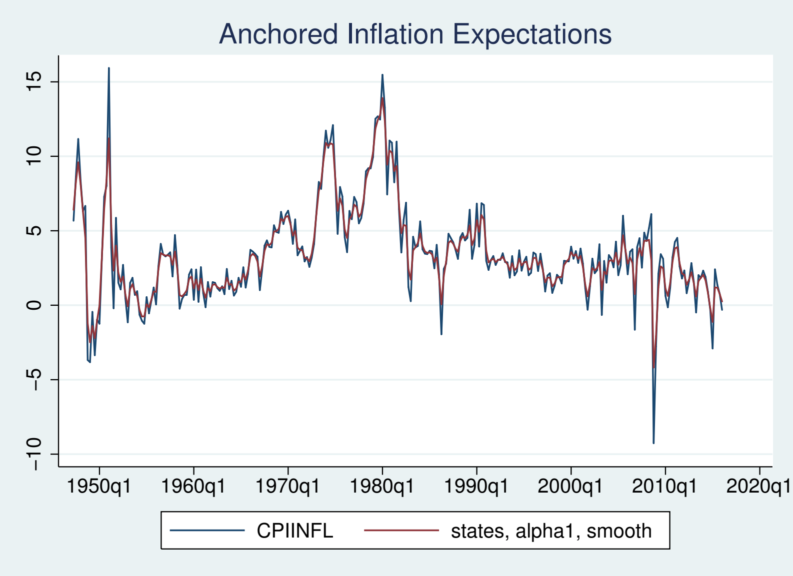
US Inflation
CPIINFL
15
10
5
0
−5
−10
1950q1
1960q1
1970q1
1980q1
1990q1
2000q1
2010q1
2020q1

Example: US Inflation (continued)

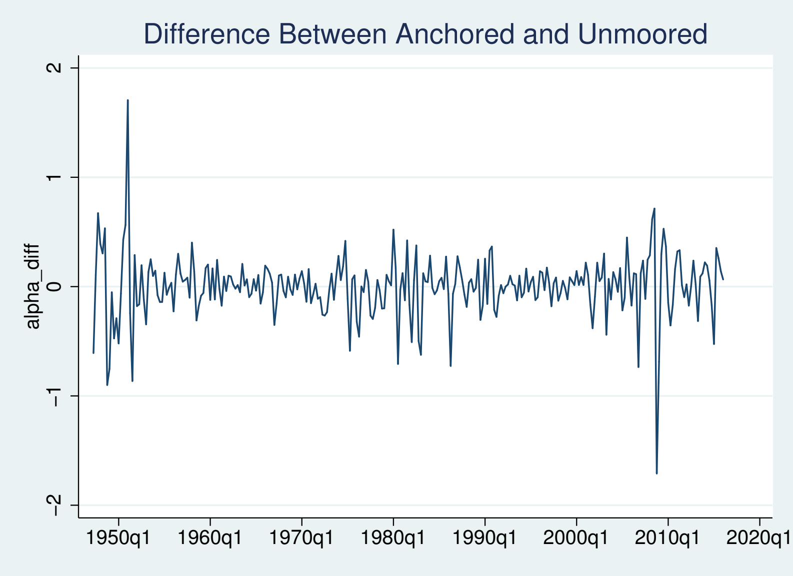
- Next figure graphs US CPI inflation against estimate of α_t using local level model
- Unmoored inflation expectations
- Following figure does same with anchored inflation expectations model
- One after that takes difference between two estimates
- Note: central bank researchers usually use models more sophisticated than this (stochastic volatility)
- Estimates of trend inflation track actual inflation too closely
- But maybe that is what you would expect if unmoored or poorly anchored



Unmoored Inflation Expectations
15
10
5
0
−5
−10
1950q1
1960q1
1970q1
1980q1
1990q1
2000q1
2010q1
2020q1
CPIINFL
states, alpha, smooth



Anchored Inflation Expectations
15
10
5
0
−5
−10
1950q1
1960q1
1970q1
1980q1
1990q1
2000q1
2010q1
2020q1
CPIINFL
states, alpha1, smooth


Difference Between Anchored and Unmoored
alpha_diff
2
1
0
−1
−2
1950q1
1960q1
1970q1
1980q1
1990q1
2000q1
2010q1
2020q1

Econometric Estimation of Local Level Model

- Remember form of local level model is:

$$y_t = \alpha_t + \varepsilon_t$$

- ε_t is i.i.d. $N\left(0, \sigma_\varepsilon^2\right)$.
- α_t is the state and follows

$$\alpha_{t+1} = \alpha_t + u_t$$

- u_t is i.i.d. $N\left(0, \sigma_u^2\right)$
- Model has two parameters to estimate: σ_ε^2 and σ_u^2
- And want estimates of the states
- Important point for empirical researcher: Computer software packages such as R do this

Filtering versus Smoothing in the Local Level Model

- Notation: superscripts for all observations up to a specific time
- E.g. $y^T = (y_1, .., y_T)'$ is all observations in the sample
- $\alpha^t = (\alpha_1, .., \alpha_t)'$ is all states up to the current period (t)
- Filtering = using y^t
- $E(\alpha_t | y^t)$ is the filtered estimate of the state
- Notation: "$|$" as in "$|y^t$" means conditional on or given
- $E(y_{t+1} | y^t)$ is estimate of y_{t+1} (unknown at time t)
- Used for real time forecasting
- Smoothing = using y^T
- $E(\alpha_t | y^T)$ is smoothed estimate of state
- E.g. estimate of trend inflation using the full sample of data

The Kalman Filter

- I will not derive or state exact formulae, just the main ideas
- Formulae below depend on σ_ε^2 and σ_u^2, for now assume known
- Can prove

$$\begin{aligned} \alpha_t | y^{t-1} &\sim N\left(a_{t|t-1}, P_{t|t-1}\right) \\ \alpha_t | y^t &\sim N\left(a_{t|t}, P_{t|t}\right) \end{aligned}$$

- First called *prediction equation*
- Second called *updating equation*
- Kalman filter involves simple formulae linking $a_{t|t-1}, P_{t|t-1}, a_{t|t}, P_{t|t}$
- Also formula for predictive density $p\left(y_{t+1}|y^t\right)$ which can be used for real time forecasting
- Also formula for likelihood function (used for maximum likelihood estimation)

Kalman Filter Recursions

- Start with initial condition, $a_{1|1}, P_{1|1}$ (to be discussed)
- Calculate $a_{2|1}$, $P_{2|1}$ using Kalman filtering formulae
- Calculate $a_{2|2}$, $P_{2|2}$
- ...
- Calculate $a_{t|t-1}$, $P_{t|t-1}$
- Calculate $a_{t|t}$, $P_{t|t}$
- etc.

Kalman Filter Recursions

- Each calculation on previous slide only depended on the last one
- New observation added, only need to update using this
- Simplifies computation: no need for manipulations involving $T \times T$ matrices
- At every point in time get filtered estimate of state, predictive density, etc.
- Run the Kalman filter from $t = 1, .., T$

Prediction Error Decomposition

- Kalman filter recursions above for states (e.g. $a_{t|t}$ and $a_{t|t-1}$ estimates of states given different information sets)
- But also same thing for forecasts: $y_{t|t-1}$
- Note: nothing for $y_{t|t}$ since $y_{t|t} = y_t$ (you already observe it)
- Can show

$$y_t|y^{t-1} \sim N\left(y_{t|t-1}, F_t\right)$$

- where $y_{t|t-1}$ and F_t have simple recursive formula ("one observation at a time")
- Prediction error:

$$v_t = y_t - y_{t|t-1}$$

- Can show likelihood function depends only on v_t and F_t
- Called prediction error decomposition

Initializing the Kalman Filter

- Need to start Kalman filter with initial condition, $a_{1|1}, P_{1|1}$
- Standard treatments exist (R does this automatically)
- But basic idea is to assume

$$\alpha_1 \sim N\left(a_{1|1}, P_{1|1}\right)$$

- To describe ideas, work with generalization of local level model

$$\alpha_{t+1} = c + \rho\alpha_t + u_t$$

- This implies

$$\begin{aligned} \alpha_1 &= c + \rho\alpha_0 + u_1 \\ &= c + \rho\left(c + \rho\alpha_{-1} + u_0\right) + u_1 \\ &= etc. \end{aligned}$$

Initializing the Kalman Filter

- Can show, if $|\rho| < 1$:

$$
\begin{aligned}
E(\alpha_1) &= \frac{c}{(1-\rho)} \\
var(\alpha_1) &= \frac{\sigma_u^2}{1-\rho^2}
\end{aligned}
$$

- Setting $a_{1|1} = E(\alpha_1)$ and $P_{1|1} = var(\alpha_1)$ is called the stationary initialization
- But only works if $|\rho| < 1$ (or, more generally, if state equation is stationary)
- For nonstationary state equations (like local level model) "diffuse initialization" done
- Amounts to letting $P_{1|1} \to \infty$ and can show $a_{1|1}$ will not appear in formula

Maximum Likelihood

- Preceding material discussed estimating states and forecasting assuming σ_ε^2 and σ_u^2, known
- How do we estimate them?
- Maximum likelihood using prediction error decomposition
- There are slightly different variants based on initialization of Kalman filter

Model Selection and Hypothesis Testing

- Likelihood based testing and model selection can be done
- Tests of whether a parameter is significantly different from zero or likelihood ratio tests
- Information criteria such as AIC and BIC can be used
- Information criteria described on pages 48-49 of Tsay textbook
- Two state space models with same y_t choose one with lower AIC or BIC

Digression: Information Criteria

- Information criteria are a popular method of model selection
- Can be used to choose between any models (not just state space models)
- Let $L(y;\theta)$ be the likelihood function for a model which depends on parameters θ
- Let p be the number of parameters in θ
- In the local level model θ will be σ_ε^2 and σ_u^2 and, thus, $p = 2$

Digression: Information Criteria

- Information criteria typically have the form:

$$IC(\theta) = -2 \ln [L(y;\theta)] + g(p) \tag{1}$$

- $g(p)$ is an increasing function of p.
- Evaluate $IC(\theta)$ at the maximum likelihood value for θ for every model under consideration
- Choose the model with the lowest IC
- Information criteria differ in the choice of of $g(p)$.
- This is a function which rewards parsimony (penalizes models with excessive parameters)

Digression: Information Criteria

- There are many ICs
- Two of the most popular ones are Bayesian Information Criterion (or BIC) and Akaike's Information Criterion (or AIC)
- BIC is sometimes called Schwarz Criterion
- BIC sets $g(p) = p \ln(T)$
- AIC sets $g(p) = 2p$

Example: US Inflation (continued)

- Illustrate using US CPI Inflation example
- Estimate local level model (parameters σ_ε^2 and σ_u^2)
- Table provided for every parameter being estimated along with t-stat and P-value for whether parameter equals zero
- For σ_u^2: z= 4.22 and P>|z|=0.000
- Strong evidence that σ_u^2 is not zero 0
- Thus α_t is varying over time (non-stationary)
- Thus, unit root is present

Example: US Inflation (continued)

- Is local level model better than stationary model?

$$\alpha_{t+1} = c + \rho\alpha_t + u_t$$

- Local level model gives BIC of 1238.123
- Stationary model gives BIC of 1235.538
- Choose stationary model

State Smoothing

- Smoothing uses full sample, y^T
- Suitable for estimation (e.g. estimating trend inflation)
- Standard recursive formulae exist with same "update one observation at a time"
- Can prove
$$\alpha_t | y^T \sim N\left(a_{t|T}, P_{t|T}\right)$$
- First run Kalman filter from $t = 1, .., T$
- Then state smoother from $t = T, .., 1$
- Set of simple recursive formulae for $a_{t|T}$ and $P_{t|T}$

Missing Values

- What if you have missing observations?
- Kalman filter and state smoother can handle this
- Kalman filter recursive calculation of $a_{t|t-1}, P_{t|t-1}, a_{t|t}, P_{t|t}$
- If observation is missing at time τ, can show no updating occurs
- $a_{\tau|\tau} = a_{\tau|\tau-1}$ and $P_{\tau|\tau} = P_{\tau|\tau-1}$
- Simple way of dealing with missing values which occur in many data sets

Structural Time Series Models

- Structural Time Series (STS) models popularized by Andrew Harvey
- Also called Unobserved Components Models
- Local level model is one example
- Many more are popular
- Which one you use depends on the properties of the data set
- I will give a few more examples

Local Linear Trend Model

- Same measurement equation as local level model:

$$y_t = \alpha_t + \varepsilon_t$$

- And α_t still interpreted as a trend, but

$$\alpha_{t+1} = \alpha_t + \beta_t + u_{1t}$$

with additional state equation:

$$\beta_{t+1} = \beta_t + u_{2t}$$

- In computer lab I ask you to estimate and plots the two states and gain more understanding
- Trend has its own trend: I(2)

$$\Delta y_t = \beta_{t-1} + \varepsilon_t - \varepsilon_{t-1} + u_{1,t-1}$$

but β_{t-1} still has unit root

Adding a Cycle

- The business cycle affects many macro variables
- Macro variables might have trend, irregular component and a cyclical component
- Simple to add cycle to local level model:

$$y_t = \alpha_t + \phi_t + \varepsilon_t$$

- α_t and ε_t same as for local level model
- ε_t is i.i.d. $N\left(0, \sigma_\varepsilon^2\right)$.
- α_t is a state and follows

$$\alpha_{t+1} = \alpha_t + u_t$$

- u_t is i.i.d. $N\left(0, \sigma_u^2\right)$

Adding a Cycle

- New state is ϕ_t used to estimate the cycle
- Need to use some trigonometry to model cycle
- Do not worry if unfamiliar: key thing is that sines and cosines are functions that can easily be evaluated
- Cycles are characterized by their frequency (e.g. how often recessions occur) and amplitude (e.g. how deep recessions are)
- Can be modeled as

$$\phi_{t+1} = a \cos(\lambda t) + b \sin(\lambda t)$$

where $t = 1, .., T$ is time

Adding a Cycle

- This state is a deterministic function of time (but still a state space model)
- *a*, *b* and (usually) λ are parameters which can be estimated
- It is common to allow the cycle to change over time (e.g. business cycles used to happen more frequency in the past then recently)
- This can also be done in a state space framework

Adding Seasonality

- Many macro variables exhibit a seasonal pattern
- E.g. retail sales peaks around Christmas
- May want to put in seasonal pattern to an STS model
- Note: data sometimes comes already seasonally adjusted then no need to do this
- Many ways of modelling seasonality: seasonal dummies, sines and cosines, seasonal differencing
- Here will illustrate one STS approach with quarterly data that works well with many data sets

Adding Seasonality

- STS with trend, irregular component and seasonality is:

$$y_t = \alpha_t + \gamma_t + \varepsilon_t$$

- γ_t is seasonal effect
- α_t is trend, ε_t irregular component
- Can add cycle if you want
- α_t is the level (mean, intercept) of the series
- γ_t will be added to this

Adding Seasonality

- Let $y_t, y_{t-1}, y_{t-2}, y_{t-3}$ be variable in four quarters in a year
- Mean of variable in last quarter is $\alpha_t + \gamma_t$
- Mean of variable in second last quarter is $\alpha_{t-1} + \gamma_{t-1}$
- etc.
- To keep interpretation α_t as the level of series restrict:

$$\gamma_t + \gamma_{t-1} + \gamma_{t-2} + \gamma_{t-3} = 0$$

- Seasonal effects constrained to sum to zero over the year
- α_t is still the trend/average mean, seasonal effects are deviations from this
- E.g. $\gamma_t = \gamma_{t-1} = -1$ and $\gamma_{t-2} = \gamma_{t-3} = 1$
- First 2 quarters of year are 1 above trend and last two quarters of year are 1 below trend

Adding Seasonality

- But perhaps this seasonal pattern evolves over time so

$$\gamma_t + \gamma_{t-1} + \gamma_{t-2} + \gamma_{t-3} = w_t$$

  where w_t is i.i.d. $N\left(0, \sigma_w^2\right)$
- Equivalently

$$\gamma_t = -\gamma_{t-1} - \gamma_{t-2} - \gamma_{t-3} + w_t$$

- This can be put in state space form
- State space estimation methods can be used

Adding Seasonality

- Let $\gamma_t, \gamma_{t-1}, \gamma_{t-2}$ be three different states (γ_{t-3} is one minus the sum of these)
- Define a vector of states: $\theta_t = (\alpha_t, \gamma_t, \gamma_{t-1}, \gamma_{t-2})'$
- Measurement equation:

$$y_t = (1\ 1\ 0\ 0)\, \theta_t + \varepsilon_t$$

- State equation

$$\theta_{t+1} = \begin{bmatrix} \alpha_{t+1} \\ \gamma_{t+1} \\ \gamma_t \\ \gamma_{t-1} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & -1 & -1 & -1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} \alpha_t \\ \gamma_t \\ \gamma_{t-1} \\ \gamma_{t-2} \end{bmatrix} + \begin{bmatrix} u_t \\ w_t \\ 0 \\ 0 \end{bmatrix}$$

Adding Seasonality

- Or

$$\theta_{t+1} = C\theta_t + v_t$$

- Where

$$C = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & -1 & -1 & -1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

- C is a matrix of constraints in state equation
- $(1\ 1\ 0\ 0)$ are constraints in measurement equation

Adding Seasonality

- Key point: STS with seasonality is a state space model and can be estimated as such
- Some of the state equations look odd
- Last two are simply identities with no error
- But this is not a problem for the Kalman filter and state space methods

Another Example: the ARMA(2,2) Model

- ARMA models are a popular class of time series models
- Here we will show how ARMA(2,2) can be written as a state space model
- Similar derivations hold for ARMA(p,q)
- ARMA(2,2) model is

$$y_t = a_1 y_{t-1} + a_2 y_{t-2} + \eta_t + b_1 \eta_{t-1} + b_2 \eta_{t-2}$$

Another Example: the ARMA(2,2) Model

- Now consider a state space model with 4 states with state equations

$$\alpha_t = \begin{bmatrix} \alpha_{1,t} \\ \alpha_{2,t} \\ \alpha_{3,t} \\ \alpha_{4,t} \end{bmatrix} = \begin{bmatrix} a_1 & a_2 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \alpha_{1,t-1} \\ \alpha_{2,t-1} \\ \alpha_{3,t-1} \\ \alpha_{4,t-1} \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \\ b_1 \\ b_2 \end{bmatrix} \eta_t$$

- And measurement equation

$$y_t = (1\ 0\ 0\ 0)\, \alpha_t$$

- Will show this state space model is equivalent to ARMA(2,2)

Another Example: the ARMA(2,2) Model

- First state equation is

$$\alpha_{1,t} = a_1\alpha_{1,t-1} + a_2\alpha_{2,t-1} + \alpha_{3,t-1} + \eta_t$$

- Substitute in formulae for $\alpha_{2,t-1}$ and $\alpha_{3,t-1}$ from second and third state equations gives

$$\alpha_{1,t} = a_1\alpha_{1,t-1} + a_2\alpha_{1,t-2} + \alpha_{4,t-2} + b_1\eta_{t-1} + \eta_t$$

- The fourth state equation implies $\alpha_{4,t-2} = b_2\eta_{t-2}$ and, thus,

$$\alpha_{1,t} = a_1\alpha_{1,t-1} + a_2\alpha_{1,t-2} + b_2\eta_{t-2} + b_1\eta_{t-1} + \eta_t$$

Another Example: the ARMA(2,2) Model

- But the measurement equation says $\alpha_{1,t} = y_t$ and thus,

$$y_{1,t} = a_1 y_{1,t-1} + a_2 y_{1,t-2} + b_2 \eta_{t-2} + b_1 \eta_{t-1} + \eta_t$$

- This is the original ARMA(2,2) model
- In general, any ARMA(p,q) model can be written as a state space model

The Normal Linear State Space Model

- All the models above are examples of Normal Linear State Space Models
- Measurement equation:

$$y_t = W_t \delta + Z_t \beta_t + \varepsilon_t$$

- State equation:

$$\beta_{t+1} = D_t \beta_t + u_t$$

- y_t contains dependent variable(s)
- Note: so far models all have y_t being one variable (e.g. inflation)
- But y_t can be a vector $M \times 1$ vector
- Usual for macroeconomics: VARs have M variables, DSGE models involve M variables, Dynamic Factor Model with many variables
- Big Data applications often have M being large (e.g. 100 or more)
- β_t is a $K \times 1$ vector of states

The Normal Linear State Space Model

- Errors in measurement and state equations both can be vectors
- Need error covariance matrices (not just variances)
- Also may want to have these varying over time (e.g. GARCH effects)
- Hence, this very general model allows for:
- $\varepsilon_t \sim N(0, \Sigma_t)$
- $u_t \sim N(0, Q_t)$.
- ε_t and u_t are independent of each other and over time

The Normal Linear State Space Model

- The remaining terms are called *system matrices*
- $W_t, Z_t, \Sigma_t, Q_t, D_t$ and δ
- Either known or want to estimate them
- W_t is $M \times p_0$ matrix
- Z_t is $M \times K$ matrix
- D_t is a $K \times K$ matrix

The Normal Linear State Space Model: Examples

- You will usually work with special cases of this very general modelling framework
- Warning: Some computer packages requires many system matrices to be constant:
- $Z_t = Z, D_t = D, \Sigma_t = \Sigma, Q_t = Q$
- All of the STS models discussed so far had $W_t = 0$
- If W_t contains explanatory variables, you combine regression $(W_t\delta)$ with state space model $(Z_t\beta_t)$ parts into one model
- Can also add explanatory variables in your state equation
- Local level model: $Z_t = D_t = 1$, $\Sigma_t = \sigma_\varepsilon^2$ and $Q_t = \sigma_u^2$
- Note: random walk behaviour of trend can be relaxed by setting $D_t = \rho$ and estimating ρ

The Normal Linear State Space Model: Examples

- STS trend plus seasonality model had $\Sigma_t = \sigma_\varepsilon^2$, $Z_t = (1\ 1\ 0\ 0)$

$$D_t = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & -1 & -1 & -1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

- $$Q_t = \begin{bmatrix} \sigma_u^2 & 0 & 0 & 0 \\ 0 & \sigma_w^2 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

DSGE Models as State Space Models

- DSGE = Dynamic, stochastic general equilibrium models popular in modern macroeconomics and commonly used in policy circles (e.g. central banks).
- I will not explain the macro theory, other than to note they are:
- Derived from microeconomic principles (based on agents and firms decision problems), dynamic (studying how economy evolves over time) and general equilibrium.
- Solution (using linear approximation methods) is a linear state space model
- Note: recent work with second order approximations yields nonlinear state space model
- Nonlinear state space models hot field now

Estimation Strategy for DSGE

- Most linearized DSGE models written as:

$$\Gamma_0(\theta)\, s_t = \Gamma_1(\theta)\, E_t(s_{t+1}) + \Gamma_2(\theta)\, s_{t-1} + \Gamma_3(\theta)\, u_t$$

- s_t is vector containing both observed variables (e.g. output growth, inflation, interest rates) and unobserved variables (e.g. technology shocks, monetary policy shocks).
- Note, theory usually written in terms of s_t as deviation of variable from steady state (an issue I will ignore here to keep exposition simple)
- θ are structural parameters (e.g. parameters for steady states, tastes, technology, policy and driving the exogenous shocks).
- u_t are structural shocks ($N(0, I)$).
- $\Gamma_j(\theta)$ are often highly nonlinear functions of θ

Solving the DSGE Model

- Methods exist to solve linear rational expectations models such as DSGE models
- If unique equilibrium exists can be written as:

$$s_t = A(\theta)\, s_{t-1} + B(\theta)\, u_t$$

- Looks like a VAR, but....
- Some elements of s_t unobserved
- Interpret this as our state equation
- and highly nonlinear restrictions involved in $A(\theta)$ and $B(\theta)$

DSGE Model as a State Space Model

- Let y_t be elements of s_t which are observed.
- Measurement equation:

$$y_t = Cs_t$$

  where C is matrix which picks out observed elements of s_t
- This is a restricted version of a state space model

DSGE Model as a State Space Model

- In terms of notation for Normal linear state space model we have:
- States: $\beta_t = s_t$
- Restrictions on how states impact on the data: $Z_t = C$
- Restrictions in state equation: $D_t = A(\theta)$
- Error covariance in state equation: $Q_t = B(\theta) B(\theta)'$
- Error covariance in measurement equation: $\Sigma_t = 0$
- Note: Standard software packages cannot handle nonlinearity in θ so must use specialized software
- E.g. DYNARE (http://www.dynare.org/)

Estimation of Normal Linear State Space Models

- I previously discussed estimation for local level model
- Exactly the same idea for Normal Linear State Space model but more general formulae
- Kalman filter and state smoother to estimate states
- Prediction error decomposition used to calculate likelihood function
- Likelihood function: use for MLE of any unknown parameters in $W_t, Z_t, \Sigma_t, Q_t, D_t$ and δ
- Likelihood based methods for testing or model selection (information criteria) can be done

Nonlinear, Non-normal State Space Models

- Techniques covered in this lecture (e.g. Kalman filter) require linearity and Normal errors
- Some interesting models are nonlinear and/or non-Normal
- E.g. Normal errors sometimes inappropriate (e.g. if variable is a count, etc.)
- E.g. DSGE models which are not linearized
- Much recent research developing econometric methods for nonlinear state space models
- One important nonlinear model is stochastic volatility

Stochastic Volatility

- In finance, popular alternative to ARCH and GARCH for modelling time-varying volatility
- Stochastic volatility model:

$$y_t = \exp\left(\frac{h_t}{2}\right)\varepsilon_t$$

-

$$h_{t+1} = h_t + \eta_t$$

- ε_t is i.i.d. $N(0,1)$ and η_t is i.i.d. $N\left(0,\sigma_\eta^2\right)$. ε_t and η_s are independent.
- h_t is log of the variance of y_t (log volatility)
- This is state space model with states being h_t, but measurement equation is not a linear function of h_t

Stochastic Volatility

- Exact maximum likelihood estimation of this nonlinear state space model can be hard (Bayesian estimation is common)
- But approximate estimation can be done easily
- Can use transformation to make approximately Normal linear state space model
- Approximation can be poor (esp. in finance)
- Square and log the measurement equation:

$$y_t^* = h_t + \varepsilon_t^*$$

- where $y_t^* = \ln\left(y_t^2\right)$ and $\varepsilon_t^* = \ln\left(\varepsilon_t^2\right)$.
- Now the measurement equation is linear so maybe we can use algorithm for Normal linear state space model?
- No, since error is no longer Normal (i.e. $\varepsilon_t^* = \ln\left(\varepsilon_t^2\right)$)
- But if you approximate the distribution of ε_t^* by Normal can use Kalman filter, etc.
- Same as local level model but with dependent variable y_t^*

Summary

- A set of econometric tools exist for estimating and forecasting with state space models
- Kalman filter, state smoother, etc.
- Many models of interest can be put in a state space framework including:
- Almost everything you covered in earlier study of time series econometrics plus:
- Structural Time Series models
- DSGE models
- Time-varying parameter models (future lecture)
- Factor models (future lecture)
- Stochastic volatility (alternative to GARCH)