7 Day 3: Time Varying Parameter Models

References:

1. Durbin, J. and S.-J. Koopman (2001). *Time Series Analysis by State Space Methods*. Oxford University Press, Oxford

2. Koopman, S.-J., N. Shephard, and J.A. Doornik (2001). "Statistical Algorithms for State Space Models Using SsfPack 2.2," *Econometrics Journal*, 2, 113-166.

3. Zivot, E. and J. Wang (2002). *Modeling Financial Time Series with S-PLUS*. Springer-Verlag, New York.

7.1 Rolling Regression

For a window of width $k < n < T$, the rolling linear regression model is

$$\underset{(n\times 1)}{\mathbf{y}_t(n)} = \underset{(n\times k)}{\mathbf{X}_t(n)}\underset{(k\times 1)}{\boldsymbol{\beta}_t(n)} + \underset{(n\times 1)}{\boldsymbol{\varepsilon}_t(n)},\ t = n, \ldots, T$$

- Observations in $y_t(n)$ and $X_t(n)$ are n most recent values from times $t - n + 1$ to t
- OLS estimates are computed for sliding windows of width n and increment m
- Poor man's time varying regression model

7.1.1 Application: Simulated Data

- compute rolling regressions for 24-month windows incremented by 1 month

7.1.2 Application: Exchange Rate Data

- compute rolling regressions for 24-month windows incremented by 1 month
- compute rolling regressions for 48-month windows incremented by 12 months

7.2 Time Varying Parameter Regression Model

References:

The most used TVP regression has the form

$$\begin{aligned} y_t &= \beta_{0,t} + \beta_{1,t}x_{1t} + \cdots + \beta_{k,t}x_{kt} + \nu_t,\ \nu_t \sim N(0, \sigma_\nu^2) \\ \beta_{i,t+1} &= \beta_{i,t} + \xi_{i,t},\ \xi_{i,t} \sim N(0, \sigma_i^2),\ i = 0, \ldots, k \end{aligned}$$

Remarks:

- Random walk specification captures variety of parameter variation
- Model is most conveniently estimated and analyzed using state space methods

7.3 Linear Gaussian State Space Models

$$
\begin{aligned}
\underset{m\times 1}{\boldsymbol{\alpha}_{t+1}} &= \underset{m\times 1}{\mathbf{d}_t} + \underset{m\times m}{\mathbf{T}_t} \cdot \underset{m\times 1}{\boldsymbol{\alpha}_t} + \underset{m\times r}{\mathbf{H}_t} \cdot \underset{r\times 1}{\boldsymbol{\eta}_t} \\
\underset{N\times 1}{\mathbf{y}_t} &= \underset{N\times 1}{\mathbf{c_t}} + \underset{N\times m}{\mathbf{Z_t}} \cdot \underset{m\times 1}{\boldsymbol{\alpha}_\mathbf{t}} + \underset{N\times N}{\mathbf{G}_t} \cdot \underset{N\times 1}{\boldsymbol{\varepsilon}_t}
\end{aligned}
$$

where $t = 1, \ldots, n$ and

$$
\begin{aligned}
&\boldsymbol{\alpha}_1 \sim N(\mathbf{a}, \mathbf{P}), \boldsymbol{\eta}_t \sim iid\ N(0, \mathbf{I}_r), \boldsymbol{\varepsilon}_t \sim iid\ N(\mathbf{0}, \mathbf{I}_N) \\
&E[\boldsymbol{\varepsilon}_t \boldsymbol{\eta}_t'] = \mathbf{0}
\end{aligned}
$$

Compact notation used by `SsfPack`

$$
\begin{aligned}
\begin{pmatrix} \boldsymbol{\alpha}_{t+1} \\ \mathbf{y}_t \end{pmatrix} &= \underset{(m+N)\times 1}{\boldsymbol{\delta}_t} + \underset{(m+N)\times m}{\boldsymbol{\Phi}_t} \cdot \underset{m\times 1}{\boldsymbol{\alpha}_t} + \underset{(m+N)\times 1}{\mathbf{u}_t}, \\
& \boldsymbol{\alpha}_1 \sim N(\mathbf{a}, \mathbf{P}) \\
& \mathbf{u}_t \sim iid\ N(\mathbf{0}, \boldsymbol{\Omega}_t)
\end{aligned}
$$

where

$$
\begin{aligned}
\boldsymbol{\delta}_t &= \begin{pmatrix} \mathbf{d}_t \\ \mathbf{c}_t \end{pmatrix}, \ \boldsymbol{\Phi}_t = \begin{pmatrix} \mathbf{T}_t \\ \mathbf{Z}_t \end{pmatrix}, \ \mathbf{u}_t = \begin{pmatrix} \mathbf{H}_t \boldsymbol{\eta}_t \\ \mathbf{G}_t \boldsymbol{\varepsilon}_t \end{pmatrix}, \\
\boldsymbol{\Omega}_t &= \begin{pmatrix} \mathbf{H}_t \mathbf{H}_t' & \mathbf{0} \\ \mathbf{0} & \mathbf{G}_t \mathbf{G}_t' \end{pmatrix}
\end{aligned}
$$

Initial value parameters

$$
\boldsymbol{\Sigma} = \begin{pmatrix} \mathbf{P} \\ \mathbf{a}' \end{pmatrix}
$$

Note: For multivariate models, i.e. $N > 1$, $\mathbf{G}_t \mathbf{G}_t'$ is assumed diagonal.

7.3.1 Initial Conditions

Initial state variance is assumed to be of the form

$$
\mathbf{P} = \mathbf{P}_* + \kappa \mathbf{P}_\infty
$$

$$
\kappa = 10^7
$$

P_* is for stationary state components

P_∞ is for non-stationary state components

7.3.2 Regression Model with Time Varying Parameters

$$\begin{aligned} y_t &= \beta_{0,t} + \beta_{1,t} x_t + \nu_t,\ \nu_t \sim N(0, \sigma_\nu^2) \\ \beta_{0,t+1} &= \beta_{0,t} + \xi_t,\ \xi_t \sim N(0, \sigma_\xi^2) \\ \beta_{1,t+1} &= \beta_{1,t} + \varsigma_t,\ \varsigma_t \sim N(0, \sigma_\varsigma^2) \end{aligned}$$

Let $\boldsymbol{\alpha}_t = (\beta_{0,t}, \beta_{1,t})'$, $\mathbf{x}_t = (1, x_t)'$, $\mathbf{H}_t = diag(\sigma_\xi, \sigma_\varsigma)'$ and $G_t = \sigma_\nu$. The state space form is

$$\begin{pmatrix} \boldsymbol{\alpha}_{t+1} \\ y_t \end{pmatrix} = \begin{pmatrix} \mathbf{I}_2 \\ \mathbf{x}_t' \end{pmatrix} \boldsymbol{\alpha}_t + \begin{pmatrix} \mathbf{H}\boldsymbol{\eta}_t \\ G\varepsilon_t \end{pmatrix}$$

and has parameters

$$\boldsymbol{\Phi}_t = \begin{pmatrix} \mathbf{I}_2 \\ \mathbf{x}_t' \end{pmatrix},\ \boldsymbol{\Omega} = \begin{pmatrix} \sigma_\xi^2 & 0 & 0 \\ 0 & \sigma_\varsigma^2 & 0 \\ 0 & 0 & \sigma_\nu^2 \end{pmatrix}$$

The initial state matrix is

$$\boldsymbol{\Sigma} = \begin{pmatrix} -1 & 0 \\ 0 & -1 \\ 0 & 0 \end{pmatrix}$$

7.3.3 Regression model with fixed parameters

The regression model with fixed regressors occurs when

$$\sigma_\xi^2 = \sigma_\varsigma^2 = 0$$

7.4 Kalman Filter and Smoother

The Kalman filter is a recursive algorithm for the evaluation of moments of the normally distributed state vector $\boldsymbol{\alpha}_{t+1}$ conditional on the observed data $\mathbf{Y}_t = (y_1, \ldots, y_t)$ and the state space model parameters. Let $\mathbf{a}_t = E[\boldsymbol{\alpha}_t|\mathbf{Y}_{t-1}]$ and $\mathbf{P}_t = var(\boldsymbol{\alpha}_t|\mathbf{Y}_{t-1})$

- The *filtering* or *updating* equations compute

$$\begin{aligned} \mathbf{a}_{t|t} &= E[\boldsymbol{\alpha}_t|\mathbf{Y}_t], \\ \mathbf{P}_{t|t} &= var(\boldsymbol{\alpha}_t|\mathbf{Y}_t), \\ \mathbf{v}_t &= \mathbf{y}_t - \mathbf{c}_t - \mathbf{Z}_t\mathbf{a}_t \text{ (prediction error)}, \\ \mathbf{F}_t &= var(\mathbf{v}_t) \text{ (prediction error variance)} \end{aligned}$$

- The *prediction* equations of the Kalman filter compute $\mathbf{a}_{t+1}$ and $\mathbf{P}_{t+1}$

The *Kalman smoothing* algorithm is a backward recursion which computes the mean and variance of specific conditional distributions based on the full data set $\mathbf{Y}_n = (y_1, \ldots, y_n)$.

- The smoothed estimates of the state vector $\boldsymbol{\alpha}_t$ and its variance matrix are denoted

$$\begin{aligned}\hat{\boldsymbol{\alpha}}_t &= \mathbf{a}_{t|n} = E[\boldsymbol{\alpha}_t|\mathbf{Y}_n] \\ \mathbf{P}_{t|n} &= var(\hat{\boldsymbol{\alpha}}_t|\mathbf{Y}_n)\end{aligned}$$

  The smoothed estimate $\hat{\boldsymbol{\alpha}}_t$ is the optimal estimate of $\boldsymbol{\alpha}_t$ using all available information $\mathbf{Y}_n$.

- The smoothed estimate of the response $\mathbf{y}_t$ and its variance are computed using

$$\begin{aligned}\hat{\mathbf{y}}_t &= \mathbf{c}_t + \mathbf{Z}_t\hat{\boldsymbol{\alpha}}_t \\ var(\hat{\mathbf{y}}_t|\mathbf{Y}_n) &= \mathbf{Z}_t var(\hat{\boldsymbol{\alpha}}_t|\mathbf{Y}_n)\mathbf{Z}_t'\end{aligned}$$

- The smoothed disturbance estimates are the estimates $\boldsymbol{\varepsilon}_t$ and $\boldsymbol{\eta}_t$ based on all available information $\mathbf{Y}_n$, and are denoted

$$\begin{aligned}\hat{\boldsymbol{\varepsilon}}_t &= \varepsilon_{t|n} = E[\boldsymbol{\varepsilon}_t|\mathbf{Y}_n] \\ \hat{\boldsymbol{\eta}}_t &= \eta_{t|n} = E[\boldsymbol{\eta}_t|\mathbf{Y}_n]\end{aligned}$$

Remarks

- Recursions are easy to code up in matrix programming languages like GAUSS, MATLAB, OX, S-PLUS, R
- SsfPack by Siem-Jan Koopman is a suite of C functions to efficiently implement the Kalman Filter and related algorithms. SsfPack has implementations in OX and S-PLUS. Eviews also implements the algorithms of SsfPack

7.5 Prediction Error Decomposition of Log-Likelihood

The *prediction error decomposition* (PED) of the log-likelihood function for the unknown parameters $\boldsymbol{\varphi}$ of a state space model is

$$\begin{aligned}\ln L(\boldsymbol{\varphi}|Y_n) &= \sum_{t=1}^{n} \ln f(\mathbf{y}_t|\mathbf{Y}_{t-1};\boldsymbol{\varphi}) \\ &= -\frac{nN}{2}\ln(2\pi) - \frac{1}{2}\sum_{t=1}^{n}\left(\ln|\mathbf{F}_t| + \mathbf{v}_t'\mathbf{F}_t^{-1}\mathbf{v}_t\right)\end{aligned}$$

where $f(\mathbf{y}_t|\mathbf{Y}_{t-1};\boldsymbol{\varphi})$ is a conditional Gaussian density implied by the state space model

- The vector of prediction errors $\mathbf{v}_t$ and prediction error variance matrices $\mathbf{F}_t$ are computed from the Kalman filter recursions.

- The state-space model parameters $\boldsymbol{\varphi}$ may be estimated by maximum likelihood using $\ln L(\boldsymbol{\varphi}|Y_n)$ computed from the PED.

Remarks

- Care must be used to ensure that $\boldsymbol{\varphi}$ is identified
- Parameter transformations are often used to simplify estimation
  - Use $\varphi = \exp(\sigma^2)$ to ensure positive variance
  - Use $\varphi = \exp(p)/(1+\exp(p))$ to ensure probabilities lie between 0 and 1
  - Exploit invariance property of MLE
  - Use "delta method" to compute asymptotic variances of un-transformed parameters

7.6 Example: Simulated Random Walk Slope Data

The estimated model assumes random intercept and slope

$$
\begin{aligned}
y_t &= \alpha_t + \beta_t x_t + \varepsilon_t \\
& \quad x_t iid\ N(0,1) \\
& \quad \varepsilon_t \sim iid\ N(0,\sigma_\varepsilon^2) \\
\alpha_t &= \alpha_{t-1} + \xi_t,\ \xi_t \sim iid\ N(0.\sigma_\xi^2) \\
\beta_t &= \beta_{t-1} + \eta_t,\ \eta_t \sim iid\ N(0.\sigma_\eta^2)
\end{aligned}
$$

The true values are

$$\sigma_\varepsilon = 0.5, \sigma_\xi = 0, \sigma_\eta = 0.1$$

7.6.1 Parameter transformations

To ensure positive variances, the log-likelihood is constructed using the parameterization

$$
\begin{aligned}
\varphi_1 &= \ln(\sigma_\xi^2) \Rightarrow \sigma_\xi^2 = \exp(\varphi_1) \\
\varphi_2 &= \ln(\sigma_\eta^2) \Rightarrow \sigma_\eta^2 = \exp(\varphi_2) \\
\varphi_3 &= \ln(\sigma_\varepsilon^2) \Rightarrow \sigma_\varepsilon^2 = \exp(\varphi_3)
\end{aligned}
$$

Note that

$$-\infty < \varphi_i < \infty$$

7.6.2 Estimation results

Estimation is performed using the SsfPack functions in S+FinMetrics. The MLE of $\boldsymbol{\varphi}$ is

MLE of TVP Model			
	Coef	Std. Error	t value
φ_1	-25.89	328.7	-0.078
φ_2	-4.045	0.499	-8.091
φ_3	-1.470	0.119	-12.36

By the invariance property of MLE, the estimates of $\boldsymbol{\sigma} = \exp(\frac{1}{2}\boldsymbol{\varphi})$ are

MLE of TVP Model			
	Coef	Std. Error	t value
σ_ξ	0.000	0.000	0.006
σ_η	0.132	0.033	4.000
σ_ε	0.479	0.029	16.81

7.6.3 Delta Method

Let $\hat{\boldsymbol{\varphi}}$ be an estimator such that

$$\sqrt{n}(\hat{\boldsymbol{\varphi}} - \boldsymbol{\varphi}) \rightarrow N(\mathbf{0}, \mathbf{V})$$

Let $g(\boldsymbol{\varphi})$ be a continuous and differentiable function, independent of n. Then

$$\begin{aligned} \sqrt{n}(g(\hat{\boldsymbol{\varphi}}) - g(\boldsymbol{\varphi})) &\rightarrow N(0, GVG') \\ G &= \frac{\partial g(\boldsymbol{\varphi})}{\partial \boldsymbol{\varphi}'} \end{aligned}$$

For example, let

$$\begin{aligned} \varphi &= (\varphi_1, \varphi_2, \varphi_3)' \\ g(\varphi) &= (\exp(\varphi_1/2), \exp(\varphi_2/2), \exp(\varphi_3/2))' \\ &= (g_1(\varphi), g_2(\varphi), g_3(\varphi))' \end{aligned}$$

Then

$$\begin{aligned} G &= \begin{pmatrix} \frac{\partial g_1(\varphi)}{\partial \varphi_1} & \frac{\partial g_1(\varphi)}{\partial \varphi_2} & \frac{\partial g_1(\varphi)}{\partial \varphi_3} \\ \frac{\partial g_2(\varphi)}{\partial \varphi_1} & \frac{\partial g_2(\varphi)}{\partial \varphi_2} & \frac{\partial g_2(\varphi)}{\partial \varphi_3} \\ \frac{\partial g_3(\varphi)}{\partial \varphi_1} & \frac{\partial g_3(\varphi)}{\partial \varphi_2} & \frac{\partial g_3(\varphi)}{\partial \varphi_3} \end{pmatrix} \\ &= \begin{pmatrix} \exp(\varphi_1/2)/2 & 0 & 0 \\ 0 & \exp(\varphi_2/2)/2 & 0 \\ 0 & 0 & \exp(\varphi_3/2)/2 \end{pmatrix} \end{aligned}$$

7.7 Example: Exchange rate data

7.7.1 TVP AR(1) model for forward discount

$$f_t - s_t = \alpha_t + \beta_t(f_{t-1} - s_{t-1}) + \varepsilon_t$$

The MLE of $\boldsymbol{\varphi}$ is

MLE of TVP Model			
	Coef	Std. Error	t value
φ_1	-5.991	0.583	-10.28
φ_2	-3.615	0.253	-14.29
φ_3	-6.255	0.519	-12.05

By the invariance property of MLE, the estimates of $\boldsymbol{\sigma} = \exp(\frac{1}{2}\boldsymbol{\varphi})$ are

MLE of TVP Model			
	Coef	Std. Error	t value
σ_ξ	0.050	0.015	3.433
σ_η	0.164	0.021	7.909
σ_ε	0.044	0.011	3.853

7.7.2 TVP regression model for differences regression

$$\Delta s_{t+1} = \alpha_t + \beta_t(f_t - s_t) + \varepsilon_t$$

The MLE of $\boldsymbol{\varphi}$ is

MLE of TVP Model			
	Coef	Std. Error	t value
φ_1	-20.07	NA	NA
φ_2	-2.159	NA	NA
φ_3	2.428	NA	NA

Note: Hessian fails to invert at MLE.

By the invariance property of MLE, the estimates of $\boldsymbol{\sigma} = \exp(\frac{1}{2}\boldsymbol{\varphi})$ are

MLE of TVP Model			
	Coef	Std. Error	t value
σ_ξ	0.000	NA	NA
σ_η	0.339	NA	NA
σ_ε	3.367	NA	NA