STATE SPACE MODEL & KALMAN FILTERING

Date: March 15, 2018

JEN-WEN LIN, PhD, CFA

STATE SPACE MODEL

REVIEW STATE-SPACE MODEL

- State-space models, originally developed by control engineers (Kalman 1960), are useful tools for expressing *dynamic systems that involve with unobserved state variables*.
- A state-space model consists of two equations:
  - *Measurement (Observation) Equation* describes the relation between observed variables (data) and unobserved state variable.
  - *Transition (State) Equation* describes the dynamics of the state variables. The transition equation has the form of a *first-order difference equation* in the state vector.

STATE SPACE MODEL (SSM)

- Observation equation

$$y_t = \boldsymbol{F}'\boldsymbol{x}_t + \varepsilon_t$$

- State equation

$$\boldsymbol{x}_t = \boldsymbol{G}\boldsymbol{x}_{t-1} + \boldsymbol{w}_t$$

  - y_t is a univariate time series for $t = 1,2,\dots,T$
  - $\boldsymbol{x}_t$ is a $p\times 1$ state vector which follows a VAR(1) process
  - $\boldsymbol{F}$ and $\boldsymbol{G}$ are $p\times 1$ and $p\times p$ coefficient matrices, respectively
  - $\varepsilon_t \sim N(0,\sigma^2)$ and $\boldsymbol{w}_t$ follows a p-variate normal distribution with mean vector equal to a $p\times 1$ zero vector and covariance matrix $\boldsymbol{W}$.

SSM FOR AR(2) MODEL

- Model

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + a_t, \qquad a_t \quad \sim NID(0, \sigma^2)$$

- Define $\boldsymbol{x}_t = \begin{bmatrix} y_t \\ y_{t-1} \end{bmatrix}$ and $\boldsymbol{w}_t = \begin{bmatrix} a_t \\ 0 \end{bmatrix}$

- State space model
  - Observation equation:

$$y_t = [1 \quad 0]\boldsymbol{x}_t$$

  - State equation:

$$\boldsymbol{x}_t = \begin{bmatrix} \phi_1 & \phi_2 \\ 1 & 0 \end{bmatrix} \boldsymbol{x}_{t-1} + \boldsymbol{w}_t$$

ALTERNATIVE SSM FOR AR(2) MODEL

- Model

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + a_t, \qquad a_t \quad \sim NID(0, \sigma^2)$$

- Define $\boldsymbol{x}_t = \begin{bmatrix} y_t \\ \phi_2 y_{t-1} \end{bmatrix}$ and $\boldsymbol{w}_t = \begin{bmatrix} a_t \\ 0 \end{bmatrix}$

- State space model
  - Observation equation:

$$y_t = [1 \quad 0]\boldsymbol{x}_t$$

  - State equation:

$$\boldsymbol{x}_t = \begin{bmatrix} \phi_1 & 1 \\ \phi_2 & 0 \end{bmatrix} \boldsymbol{x}_{t-1} + \boldsymbol{w}_t$$

SSM FOR ARMA(1,1) MODEL

- Model

$$y_t = \phi y_{t-1} + a_t + \theta a_{t-1}, \qquad a_t \quad \sim NID(0, \sigma^2)$$

- Define $\boldsymbol{x}_t = \begin{bmatrix} y_t \\ \theta a_t \end{bmatrix}$ and $\boldsymbol{w}_t = \begin{bmatrix} a_t \\ \theta a_t \end{bmatrix}$

- State space model
  - Observation equation:

$$y_t = \begin{bmatrix} 1 & 0 \end{bmatrix} \boldsymbol{x}_t$$

  - State equation:

$$\boldsymbol{x}_t = \begin{bmatrix} \phi & 1 \\ 0 & 0 \end{bmatrix} \boldsymbol{x}_{t-1} + \boldsymbol{w}_t$$

COMBINE TWO SSM

- Two state space models may be combined

$$y_t = \boldsymbol{F}_1'\boldsymbol{x}_t + \boldsymbol{F}_2'\boldsymbol{z}_t + \varepsilon_t$$
$$\boldsymbol{x}_t = \boldsymbol{G}_1\boldsymbol{x}_{t-1} + \boldsymbol{w}_t$$
$$\boldsymbol{z}_t = \boldsymbol{G}_2\boldsymbol{z}_{t-1} + \boldsymbol{u}_t$$

- Combined SSM

$$y_t = \begin{bmatrix}\boldsymbol{F}_1' & \boldsymbol{F}_2'\end{bmatrix}\begin{bmatrix}\boldsymbol{x}_t \\ \boldsymbol{z}_t\end{bmatrix} + \varepsilon_t$$

$$\begin{bmatrix}\boldsymbol{x}_t \\ \boldsymbol{z}_t\end{bmatrix} = \begin{bmatrix}\boldsymbol{G}_1 & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{G}_2\end{bmatrix}\begin{bmatrix}\boldsymbol{x}_{t-1} \\ \boldsymbol{z}_{t-1}\end{bmatrix} + \begin{bmatrix}\boldsymbol{w}_t \\ \boldsymbol{u}_t\end{bmatrix}$$

SSM FOR REGRESSION MODEL WITH AR(2) ERROR

- Model

$$y_t = \alpha + \beta f_t + n_t, \quad n_t = \phi_1 n_{t-1} + \phi_2 n_{t-2} + a_t, \quad a_t \sim NID(0, \sigma^2)$$

- Regression model

$$y_t = [1 \quad f_t]\boldsymbol{x}_t + n_t \text{ with } \boldsymbol{x}_t = \begin{bmatrix} \alpha \\ \beta \end{bmatrix} \text{ and } \boldsymbol{x}_t = \boldsymbol{x}_{t-1}$$

- AR(2) model

$$n_t = [1 \quad 0]\boldsymbol{z}_t \text{ with } \boldsymbol{z}_t = \begin{bmatrix} n_t \\ \phi_2 n_{t-1} \end{bmatrix}$$

$$\boldsymbol{z}_t = \begin{bmatrix} \phi_1 & 1 \\ \phi_2 & 0 \end{bmatrix} \boldsymbol{z}_{t-1} + \boldsymbol{w}_t \text{ with } \boldsymbol{w}_t = \begin{bmatrix} a_t \\ 0 \end{bmatrix}$$

COMBINED SSM: REGRESSION WITH AR(2) ERRORS

- Model

$$y_t = \alpha + \beta f_t + n_t, \quad n_t = \phi_1 n_{t-1} + \phi_2 n_{t-2} + a_t, \quad a_t \sim NID(0, \sigma^2)$$

- Observation equation

$$y_t = [1, f_t, 1,0]\boldsymbol{x}_t \text{ with } \boldsymbol{x}_t = [\alpha, \beta, n_t, \phi_2 n_{t-1}]'$$

- State equation

$$\boldsymbol{x}_t = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \phi_1 & 1 \\ 0 & 0 & \phi_2 & 0 \end{bmatrix} \boldsymbol{x}_{t-1} + \begin{bmatrix} 0 \\ 0 \\ a_t \\ 0 \end{bmatrix}$$

KALMAN FILTER

Once a dynamic time series model is written as the state-space form, the Kalman filter is readily available for inference on the unobserved state vector conditional on the parameters of the model and the appropriate information set.

REVIEW STATE-SPACE MODEL

- Measurement equation

$$y_t = H_t\beta_t + \boldsymbol{Az}_t + a_t, \qquad a_t \sim NID(0, \sigma^2),$$

- Transition equation

$$\beta_t = \mu + F\beta_{t-1} + v_t, \qquad v_t \sim NID(0, Q)$$
$$E(a_t v_s') = 0, \qquad \forall t, s$$

where y_t denotes the observation of interest, β_t is a $k \times 1$ vector of state variable, H_t is an $1 \times k$ coefficient vector, F is a $k \times k$ coefficient matrix, and μ is a $k \times 1$ intercept vector. (z_t denotes exogenous variables.)

NOTATION

Notation	Meaning
ψ_t	Information set up to time t
$\beta_{t\|t-1} = E[\beta_t \| \psi_{t-1}]$	Expectation (estimate) of β_t conditional on information up to $t-1$
$P_{t\|t-1} = E\left[(\beta_t - \beta_{t\|t-1})(\beta_t - \beta_{t\|t-1})'\right]$	Covariance matrix of β_t conditional on information up to $t-1$
$\beta_{t\|t}$	Expectation (estimate) of β_t conditional on information up to t
$P_{t\|t} = E\left[(\beta_t - \beta_{t\|t})(\beta_t - \beta_{t\|t})'\right]$	Covariance matrix of β_t conditional on information up to t
$y_{t\|t-1} = E[y_t \| \psi_{t-1}] = x_t \beta_{t\|t-1}$	Forecast of y_t given information up to time $t-1$
$\eta_{t\|t-1} = y_t - y_{t\|t-1}$	Prediction error
$f_{t\|t-1} = E[\eta^2_{t\|t-1}]$	Conditional variance of the prediction error
$\beta_{t\|T} = E[\beta_t \| \psi_T]$	Expectation (estimate) of β_t conditional on information up to T
$P_{t\|T} = E\left[(\beta_t - \beta_{t\|T})(\beta_t - \beta_{t\|T})'\right]$	Covariance matrix of β_t conditional on information up to T

BASIC KALMAN FILTERING

- Prediction: At the beginning of time t, we may want to form an optimal predictor of y_t based on all the available information up to time $t-1$, i.e. $y_{t|t-1}$. Note that we need to know $\beta_{t|t-1}$ to predict $y_{t|t-1}$.
- Updating: Once y_t is realized at the end of time t, the prediction error can be calculated as $\eta_{t|t-1} = y_t - y_{t|t-1}$. This prediction error contains new information about β_t beyond that contained in $\beta_{t|t-1}$. Thus, after observing y_t, a more accurate inference can be made of β_t.
- $\beta_{t|t}$, an inference of β_t based on information up to time t, may be of the following form:

$$\beta_{t|t} = \beta_{t|t-1} + K_t \eta_{t|t-1},$$

where K_t is the weight assigned to the new information about β_t contained in the prediction error.

Figure 1 Flowchart for the basic Kalman filter

$$\beta_{0|0}, P_{0|0}, l(\theta) = 0$$

$$\beta_{t|t-1} = \tilde{\mu} + F\beta_{0|0}$$
$$P_{t|t-1} = FP_{t-1|t-1}F' + Q$$

$$\eta_{t|t-1} = y_t - y_{t|t-1} = y_t - H_t\beta_{t|t-1} - Az_t$$
$$f_{t|t-1} = H_tP_{t|t-1}H_t' + R$$

$$l(\theta) = l(\theta) - \frac{1}{2}\ln\left((2\pi)^n\left|f_{t|t-1}\right|\right) - \frac{1}{2}\eta_{t|t-1}'f_{t|t-1}^{-1}\eta_{t|t-1}$$

$$\beta_{t|t} = \beta_{t|t-1} + P_{t|t-1}H_t'f_{t|t-1}^{-1}\eta_{t|t-1}$$
$$P_{t|t} = P_{t|t-1} - P_{t|t-1}H_t'f_{t|t-1}^{-1}H_tP_{t|t-1}$$

Set $t = t - 1$

$t = 1,2, \dots, T$

$$l(\theta) = -\frac{1}{2}\sum \ln\left((2\pi)^n\left|f_{t|t-1}\right|\right) - \frac{1}{2}\sum \eta_{t|t-1}'f_{t|t-1}^{-1}\eta_{t|t-1}$$

MLE ESTIMATION AND KALMAN FILTER

For given parameters of the models, the Kalman filter provides us with prediction errors $\eta_{t|t-1}$ and its variance $f_{t|t-1}$. If $\beta_{0|0}$ and $\{e_t, v_t\}_{t=1}^{T}$ are Gaussian, the distribution of y_t conditional on ψ_{t-1} is also Gaussian, i.e.

$$y_t|\psi_{t-1} \sim N\left(y_{t|t-1}, f_{t|t-1}\right).$$

The sample log likelihood function is given by

$$\ln l(\theta) = -\frac{1}{2}\sum_{t=1}^{T} \ln\left(2\pi f_{t|t-1}\right) - \frac{1}{2}\sum_{t=1}^{T} \eta'_{t|t-1} f_{t|t-1}^{-1} \eta_{t|t-1}.$$

KALMAN SMOOTHING

- Smoothing $\beta_{t|T}$ provides us with a more accurate inference on β_t, since it uses more information than the basic filter. The following two equations can be iterated backwards for $t = T-1, T-2, \dots, 1$, to get the smoothed estimates.
- Smoothing recursion:

$$\beta_{t|T} = \beta_{t|t} + P_{t|t}F'P_{t+1|t}^{-1}\left(\beta_{t+1|T} - F\beta_{t|t} - \tilde{\mu}\right),$$

$$P_{t|T} = P_{t|t} + P_{t|t}F'P_{t+1|t}^{-1}\left(P_{t+1|T} - P_{t+1|t}\right)\left(P_{t+1|t}^{-1}\right)'FP_{t|t}',$$

where the initial values for smoothing $\beta_{T|T}$ and $P_{T|T}$ can be obtained from the last step of the basic filter.

KALMAN FILTERING EXAMPLE

- We demonstrate Kalman filtering using a time series regression model with time-varying coefficients.
- An financial application of this model to long-horizon return predictability will be demonstrated at the end of this slides.

$$y_t = x_t \beta_t + e_t, \qquad \beta_t = \tilde{\mu} + F\beta_{t-1} + v_t$$
$$e_t \sim NID(0, R), \qquad v_t \sim NID(0, Q)$$

REGRESSION WITH TIME-VARYING COEFFICIENTS

$$y_t = \sum_{i=1}^{k} \beta_{it} x_{it} + a_t, \qquad a_t \sim NID(0, \sigma^2)$$

$$\beta_{it} - \delta_i = \phi_i(\beta_{i,t-1} - \delta_i) + v_{it}, \qquad v_{it} \sim NID(0, \sigma_i^2)$$

$$E(a_t v_{is}) = 0, \forall\ t, s, i$$

Measurement Equation

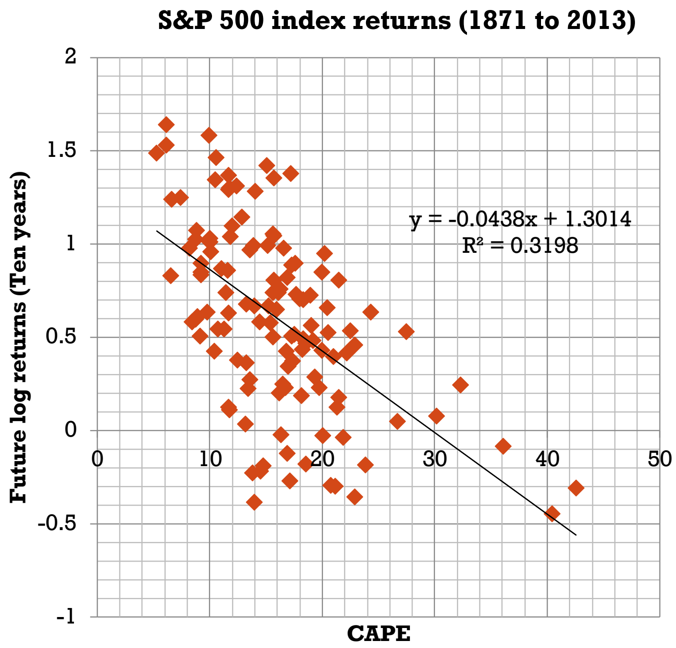
$$y_t = [x_{1t} \quad x_{2t} \quad \cdots \quad x_{kt}] \begin{bmatrix} \beta_{1t} \\ \beta_{2t} \\ \vdots \\ \beta_{kt} \end{bmatrix} + a_t$$

Transition Equation

$$\begin{bmatrix} \beta_{1t} \\ \beta_{2t} \\ \vdots \\ \beta_{kt} \end{bmatrix} = \begin{bmatrix} \delta_1^\star \\ \delta_2^\star \\ \vdots \\ \delta_k^\star \end{bmatrix} + \begin{bmatrix} \phi_1 & 0 & \cdots & 0 \\ 0 & \phi_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \phi_k \end{bmatrix} \begin{bmatrix} \beta_{1,t-1} \\ \beta_{2,t-2} \\ \vdots \\ \beta_{k,t-k} \end{bmatrix} + \begin{bmatrix} v_{1t} \\ v_{2t} \\ \vdots \\ v_{kt} \end{bmatrix}$$

where $\delta_i^\star = \delta_i(1 - \phi_i), i = 1,2, \ldots, k$.

LONG-HORIZON (LH) PREDICTABILITY OF CAMPBELL AND SHILLER (1998, JPM)



Source: http://www.econ.yale.edu/~shiller/data.htm.

- When a valuation ratio* (e.g. price to earnings ratio) is at an extreme level, either the numerator or the denominator of the valuation ratio must move in a direction that restores the ratio to a more normal level.
- Campbell and Shiller (1998) conclude that it is the stock price (denominator) that has moved to restore the ratio to its mean level.
- Campbell and Shiller (1998) tested this hypothesis by regressing forward ten year returns against Cyclically Adjusted PE ratio (CAPE).

ASSET RETURNS AND DIVIDEND DISCOUNT MODEL

- Let's start with a simple Dividend Discount Model (DDM):

$$P_0 = \frac{D_1}{r - g}. \qquad (1)$$

- Rearrange eqn. (1) as

$$\begin{aligned} \frac{E(P_1)}{P_0} - 1 = r &= \frac{D_1}{P_0} + g \\ &= \frac{D_0 \cdot g}{P_0} + g \\ &= \frac{k \cdot E_0 \cdot g}{P_0} + g. \qquad (2) \end{aligned}$$

- From DDM, we know that there is a relationship between the expected forward returns and trailing dividend yield or PE ratios.

DYNAMIC REGRESSION AND DDM

- DDM is derived under some strong assumptions so the linear relationship between expected forward returns and financial ratios (e.g. eqn. (2)) will not be held exactly in practice.
- In fact, the relationship between forward returns and financial ratios may be better described as

$$\Delta P = f(P_0, x_0, Z) + \xi$$

where $f(\ldots)$ is an unknown functional, ΔP denotes the realized forward returns, and Z stands for relevant variables that is neglected in DDM.

DYNAMIC REGRESSION AND DDM

- For any given P_0 and Z, we can approximate $f(P_0, x_0, Z)$ locally around the neighborhood of the long run valuation ratio ($\bar{x}$) using the first order Taylor approximation.
- The approximation is given by

$$\Delta P \approx f(P_0, \bar{x}, Z) + \frac{\partial f}{\partial x}(x - \bar{x}) + \xi.$$

  - Alternatively, we can express the above equation as

$$\Delta P \approx a + b(x - \bar{x}) + \xi, \qquad (4)$$

where $b = \partial f / \partial x \,|_{P_0, \bar{x}, Z}$ and $a = f(P_0, \bar{x}, Z)$.

REGRESSION WITH TIME-VARYING COEFFICIENTS (ADAPTIVE REGRESSION)

- When studying LH predictability, most researchers consider a static regression.
- Alternatively, we may consider a adaptive regression

$$y_t = a_t + b_t x_t + \varepsilon_t, \qquad (5)$$

where $a_t = a_{t-1} + w_{1t}$ and $b_t = b_{t-1} + w_{2t}$, y_t and x_t denote the forward long-horizon returns and trailing valuation ratios, respectively, and w_{it}, $i = 1{,}2$ denote noises. This model adjusts regression parameters adaptively (dynamically) according to one-step forecast errors.

- In what follows, we apply adaptive regression on LH predictability of private real estate returns.

LH PREDICTABILITY ON REAL ESTATE

Figure 1: Visualization of the future returns and predict variable

1978Q1 1978Q4 1984Q4

Averaged cap rates | Geometric averaged (annual) return

# observation	Year	Avg. cap rate	Future returns
1	1979	1978Q1-1978Q4	1979Q1-1983Q4
2	1980	1979Q1-1979Q4	1980Q1-1984Q4
3	1981	1980Q1-1980Q4	1981Q1-1985Q4
…	…	…	…
32	2010	2009Q1-2009Q4	2010Q1-2014Q4
1	2011	2010Q1-2010Q4	2011Q1-2015Q4
2	2012	2011Q1-2011Q4	2012Q1-2016Q4
3	2013	2012Q1-2012Q4	2013Q1-2017Q4
4	2014	2013Q1-2013Q4	2014Q1-2018Q4
5	2015	2014Q1-2014Q4	2015Q1-2019Q4

UPDATING FLOW OF KALMAN RECURSION
ILLUSTRATION OF ADAPTIVE REGRESSION

Time	0		1		2		3		4
Adaptive regression parameters	b0		b1		b2		b3		...
(prior / posterior)	a0		a1		a2		a3		...
Known input	x1		x2		x3		x4		...
one-step ahead forecast	yhat1		yhat2		yhat3		yhat4		...
observation	y1		y2		y3		y4		...
forecast error	y1-yhat1		y2-yhat2		y3-yhat3		y4-yhat4		...

Note that yhati denotes the one-step ahead forecast at time i, i.e. $y_{i|i-1}$. Thus, yi-yhati denotes the forecast error at time i.

LH PREDICTABILITY USING ADAPTIVE REGRESSION

PRIVATE REAL ESTATE (OFFICE SECTOR)