

State Space Models

Defn: A *state space model* for an N -dimensional time series \mathbf{y}_t consists of a *measurement equation* relating the observed data to an m -dimensional state vector $\boldsymbol{\alpha}_t$, and a Markovian transition equation that describes the evolution of the state vector over time.

The *measurement equation* has the form

$$\begin{aligned} \mathbf{y}_t &= \mathbf{Z}_t \boldsymbol{\alpha}_t + \mathbf{d}_t + \boldsymbol{\varepsilon}_t, \quad t = 1, \dots, T \\ N \times 1 & \quad (N \times m)(m \times 1) \quad N \times 1 \quad N \times 1 \\ \boldsymbol{\varepsilon}_t &\sim \text{iid } N(\mathbf{0}, \mathbf{H}_t) \end{aligned}$$

The *transition equation* for the state vector $\boldsymbol{\alpha}_t$ is the first order Markov process

$$\begin{aligned} \boldsymbol{\alpha}_t &= \mathbf{T}_t \boldsymbol{\alpha}_{t-1} + \mathbf{c}_t + \mathbf{R}_t \boldsymbol{\eta}_t \quad t = 1, \dots, T \\ m \times 1 & \quad (m \times m)(m \times 1) \quad m \times 1 \quad (m \times g)(g \times 1) \\ \boldsymbol{\eta}_t &\sim \text{iid } N(\mathbf{0}, \mathbf{Q}_t) \end{aligned}$$

For most applications, it is assumed that the measurement equation errors ε_t are independent of the transition equation errors

$$E[\varepsilon_t \boldsymbol{\eta}'_s] = \mathbf{0} \text{ for all } s, t = 1, \dots, T$$

The state space representation is completed by specifying the behavior of the initial state

$$\boldsymbol{\alpha}_0 \sim N(\mathbf{a}_0, \mathbf{P}_0)$$
$$E[\varepsilon_t \mathbf{a}'_0] = \mathbf{0}, E[\boldsymbol{\eta}_t \mathbf{a}'_0] = \mathbf{0} \text{ for } t = 1, \dots, T$$

The matrices $\mathbf{Z}_t, \mathbf{d}_t, \mathbf{H}_t, \mathbf{T}_t, \mathbf{c}_t, \mathbf{R}_t$ and \mathbf{Q}_t are called the *system matrices*, and contain non-random elements. If these matrices do not depend deterministically on t the state space system is called *time invariant*.

Note: If \mathbf{y}_t is covariance stationary, then the state space system will be time invariant.

Initial State Distribution for Covariance Stationary Models

If the state space model is covariance stationary, then the state vector α_t is covariance stationary. The unconditional mean of α_t , \mathbf{a}_0 , may be determined using

$$E[\alpha_t] = \mathbf{T}E[\alpha_{t-1}] + \mathbf{c} = \mathbf{T}E[\alpha_t] + \mathbf{c}$$

Solving for $E[\alpha_t]$, assuming \mathbf{T} is invertible, gives

$$\mathbf{a}_0 = E[\alpha_t] = (\mathbf{I}_m - \mathbf{T})^{-1}\mathbf{c}$$

Similarly, $\text{var}(\alpha_0)$ may be determined using

$$\begin{aligned}\mathbf{P}_0 &= \text{var}(\alpha_t) = \mathbf{T}\text{var}(\alpha_t)\mathbf{T}' + \mathbf{R}\text{var}(\eta_t)\mathbf{R}' \\ &= \mathbf{T}\mathbf{P}_0\mathbf{T}' + \mathbf{R}\mathbf{Q}\mathbf{R}'\end{aligned}$$

Then,

$$\begin{aligned}\text{vec}(\mathbf{P}_0) &= \text{vec}(\mathbf{T}\mathbf{P}_0\mathbf{T}') + \text{vec}(\mathbf{R}\mathbf{Q}\mathbf{R}') \\ &= (\mathbf{T} \otimes \mathbf{T})\text{vec}(\mathbf{P}_0) + \text{vec}(\mathbf{R}\mathbf{Q}\mathbf{R}')\end{aligned}$$

which implies that

$$\text{vec}(\mathbf{P}_0) = (\mathbf{I}_{m^2} - \mathbf{T} \otimes \mathbf{T})^{-1}\text{vec}(\mathbf{R}\mathbf{Q}\mathbf{R}')$$

Example: Stationary AR(2) model

$$y_t = \alpha + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \eta_t$$
$$\eta_t \sim \text{iid } N(0, \sigma^2)$$

Define $\alpha_t = (y_t, y_{t-1})'$ so that the transition equation becomes

$$\begin{pmatrix} y_t \\ y_{t-1} \end{pmatrix} = \begin{pmatrix} \phi_1 & \phi_2 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} y_{t-1} \\ y_{t-2} \end{pmatrix} + \begin{pmatrix} \alpha \\ 0 \end{pmatrix} + \begin{pmatrix} 1 \\ 0 \end{pmatrix} \eta_t$$

The transition equation system matrices are

$$\mathbf{T} = \begin{pmatrix} \phi_1 & \phi_2 \\ 1 & 0 \end{pmatrix}, \mathbf{R} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \mathbf{c} = \begin{pmatrix} \alpha \\ 0 \end{pmatrix}, Q = \sigma^2$$

The measurement equation is

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

which implies that

$$\mathbf{Z}_t = (1, 0), d_t = 0, \varepsilon_t = 0, H_t = 0$$

Distribution of initial state

$$\boldsymbol{\alpha}_0 \sim N(\mathbf{a}_0, \mathbf{P}_0)$$

Since $\boldsymbol{\alpha}_t = (y_t, y_{t-1})'$ is stationary

$$\begin{aligned} \mathbf{a}_0 &= E[\boldsymbol{\alpha}_t] = (\mathbf{I}_2 - \mathbf{T})^{-1} \mathbf{c} \\ &= \begin{pmatrix} 1 - \phi_1 & -\phi_2 \\ -1 & 1 \end{pmatrix}^{-1} \begin{pmatrix} \alpha \\ 0 \end{pmatrix} \\ &= \begin{pmatrix} \alpha / (1 - \phi_1 - \phi_2) \\ \alpha / (1 - \phi_1 - \phi_2) \end{pmatrix} \end{aligned}$$

For the state variance, solve

$$\text{vec}(\mathbf{P}_0) = (\mathbf{I}_4 - \mathbf{T} \otimes \mathbf{T})^{-1} \text{vec}(\mathbf{RQR}')$$

Simple algebra gives

$$\mathbf{I}_4 - \mathbf{T} \otimes \mathbf{T} = \begin{pmatrix} 1 - \phi_2^2 & -\phi_1\phi_2 & -\phi_1\phi_2 & -\phi_2^2 \\ -\phi_1 & 1 & -\phi_2 & 0 \\ -\phi_1 & -\phi_2 & 1 & 0 \\ -1 & 0 & 0 & 1 \end{pmatrix}$$
$$\text{vec}(\mathbf{RQR}') = \begin{pmatrix} \sigma^2 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

and so

$$\text{vec}(\mathbf{P}_0) = \begin{pmatrix} 1 - \phi_2^2 & -\phi_1\phi_2 & -\phi_1\phi_2 & -\phi_2^2 \\ -\phi_1 & 1 & -\phi_2 & 0 \\ -\phi_1 & -\phi_2 & 1 & 0 \\ -1 & 0 & 0 & 1 \end{pmatrix}^{-1} \begin{pmatrix} \sigma^2 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

Example: AR(2) model again

Another state space representation of the AR(2) is

$$\begin{aligned}y_t &= \mu + c_t \\c_t &= \phi_1 c_{t-1} + \phi_2 c_{t-2} + \eta_t\end{aligned}$$

The state vector is $\alpha_t = (c_t, c_{t-1})'$, which is unobservable, and the transition equation is

$$\begin{pmatrix} c_t \\ c_{t-1} \end{pmatrix} = \begin{pmatrix} \phi_1 & \phi_2 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} c_{t-1} \\ c_{t-2} \end{pmatrix} + \begin{pmatrix} 1 \\ 0 \end{pmatrix} \eta_t$$

This representation has measurement equation matrices

$$\begin{aligned}\mathbf{Z}_t &= (1, 0), \quad d_t = \mu, \quad \varepsilon_t = 0, \quad H_t = 0 \\ \mu &= \alpha / (1 - \phi_1 - \phi_2)\end{aligned}$$

The initial state vector has mean zero, and the initial covariance matrix is the same as that derived above.

Example: AR(2) model yet again

Yet another state space representation of the AR(2) model is

$$\begin{aligned}y_t &= (1 \ 0) \alpha_t \\ \alpha_t &= \begin{pmatrix} y_t \\ \phi_2 y_{t-1} \end{pmatrix} = \begin{pmatrix} \phi_1 & 1 \\ \phi_2 & 0 \end{pmatrix} \begin{pmatrix} y_{t-1} \\ \phi_2 y_{t-2} \end{pmatrix} \\ &\quad + \begin{pmatrix} \alpha \\ 0 \end{pmatrix} + \begin{pmatrix} 1 \\ 0 \end{pmatrix} \eta_t\end{aligned}$$

General Result: There may be many ways to write a model in state space form!

Example: MA(1) model

The MA(1) model

$$y_t = \mu + \eta_t + \theta\eta_{t-1}$$

Define $\alpha_t = (y_t - \mu, \theta\eta_t)$ and write

$$\begin{aligned} y_t &= (1 \ 0) \alpha_t + \mu \\ \alpha_t &= \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \alpha_{t-1} + \begin{pmatrix} 1 \\ \theta \end{pmatrix} \eta_t \end{aligned}$$

The first element of α_t is then $\theta\eta_{t-1} + \eta_t$ which is indeed $y_t - \mu$.

Example: ARMA(1,1) model

$$y_t = \mu + \phi(y_{t-1} - \mu) + \eta_t + \theta\eta_{t-1}$$

can be put in a state space form similar to the state space form for the MA(1). Define $\alpha_t = (y_t - \mu, \theta\eta_t)$ and write

$$\begin{aligned} y_t &= (1 \ 0) \alpha_t + \mu \\ \alpha_t &= \begin{pmatrix} \phi & 1 \\ 0 & 0 \end{pmatrix} \alpha_{t-1} + \begin{pmatrix} 1 \\ \theta \end{pmatrix} \eta_t \end{aligned}$$

The first element of α_t is then $\phi(y_{t-1} - \mu) + \theta\eta_{t-1} + \eta_t$ which is indeed $y_t - \mu$.

Example: ARMA(p, q) model

$$y_t = \phi y_{t-1} + \cdots + \phi_p y_{t-p} + \eta_t + \theta_1 \eta_{t-1} + \cdots + \theta_q \eta_{t-q}$$

may be put in state space form in the following way. Let $m = \max(p, q + 1)$ and re-write the ARMA(p, q) model as

$$y_t = \phi y_{t-1} + \cdots + \phi_p y_{t-m} + \eta_t + \theta_1 \eta_{t-1} + \cdots + \theta_{m-1} \eta_{t-m+1}$$

where some of the AR or MA coefficients will be zero unless $p = q + 1$. Define

$$\alpha_t = \begin{pmatrix} y_t \\ \phi_2 y_{t-1} + \cdots + \phi_p y_{t-m+1} + \theta_1 \eta_t + \cdots + \theta_{m-1} \eta_{t-m+2} \\ \vdots \\ \phi_m y_{t-1} + \theta_m \eta_t \end{pmatrix}$$

Then

$$y_t = (1 \quad \mathbf{0}'_{m-1}) \boldsymbol{\alpha}_t$$
$$\boldsymbol{\alpha}_t = \begin{pmatrix} \phi_1 & 1 & 0 & \cdots & 0 \\ \phi_2 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \phi_{m-1} & 0 & 0 & \cdots & 1 \\ \phi_m & 0 & 0 & \cdots & 0 \end{pmatrix} \boldsymbol{\alpha}_{t-1} + \begin{pmatrix} 1 \\ \theta_1 \\ \vdots \\ \theta_{m-2} \\ \theta_{m-1} \end{pmatrix} \eta_t$$

The Kalman Filter

The *Kalman filter* is a set of recursion equations for determining the optimal estimates of the state vector α_t given information available at time t , I_t . The filter consists of two sets of equations:

1. *Prediction equations*
2. *Updating equations*

To describe the filter, let

$$\mathbf{a}_t = E[\alpha_t | I_t] = \text{optimal estimator of } \alpha_t \text{ based on } I_t$$
$$\mathbf{P}_t = E[(\alpha_t - \mathbf{a}_t)(\alpha_t - \mathbf{a}_t)' | I_t] = \text{MSE matrix of } \mathbf{a}_t$$

Prediction Equations

Given \mathbf{a}_{t-1} and \mathbf{P}_{t-1} at time $t-1$, the optimal predictor of α_t and its associated MSE matrix are

$$\begin{aligned}\mathbf{a}_{t|t-1} &= E[\alpha_t | I_{t-1}] = \mathbf{T}_t \mathbf{a}_{t-1} + \mathbf{c}_t \\ \mathbf{P}_{t|t-1} &= E[(\alpha_t - \mathbf{a}_{t|t-1})(\alpha_t - \mathbf{a}_{t|t-1})' | I_{t-1}] \\ &= \mathbf{T}_t \mathbf{P}_{t-1} \mathbf{T}'_{t-1} + \mathbf{R}_t \mathbf{Q}_t \mathbf{R}'_t\end{aligned}$$

The corresponding optimal predictor of \mathbf{y}_t give information at $t-1$ is

$$\mathbf{y}_{t|t-1} = \mathbf{Z}_t \mathbf{a}_{t|t-1} + \mathbf{d}_t$$

The *prediction error* and its MSE matrix are

$$\begin{aligned}\mathbf{v}_t &= \mathbf{y}_t - \mathbf{y}_{t|t-1} = \mathbf{y}_t - \mathbf{Z}_t \mathbf{a}_{t|t-1} - \mathbf{d}_t \\ &= \mathbf{Z}_t (\alpha_t - \mathbf{a}_{t|t-1}) + \boldsymbol{\varepsilon}_t \\ E[\mathbf{v}_t \mathbf{v}'_t] &= \mathbf{F}_t = \mathbf{Z}_t \mathbf{P}_{t|t-1} \mathbf{Z}'_t + \mathbf{H}_t\end{aligned}$$

These are the components that are required to form the prediction error decomposition of the log-likelihood function.

Updating Equations

When new observations \mathbf{y}_t become available, the optimal predictor $\mathbf{a}_{t|t-1}$ and its MSE matrix are updated using

$$\mathbf{a}_t = \mathbf{a}_{t|t-1} + \mathbf{P}_{t|t-1} \mathbf{Z}'_t \mathbf{F}_t^{-1} (\mathbf{y}_t - \mathbf{Z}_t \mathbf{a}_{t|t-1} - \mathbf{d}_t)$$

$$= \mathbf{a}_{t|t-1} + \mathbf{P}_{t|t-1} \mathbf{Z}'_t \mathbf{F}_t^{-1} \mathbf{v}_t$$

$$\mathbf{P}_t = \mathbf{P}_{t|t-1} - \mathbf{P}_{t|t-1} \mathbf{Z}_t \mathbf{F}_t^{-1} \mathbf{Z}_t \mathbf{P}_{t|t-1}$$

The value \mathbf{a}_t is referred to as the *filtered estimate* of α_t and \mathbf{P}_t is the MSE matrix of this estimate. It is the optimal estimate of α_t given information available at time t .

Prediction Error Decomposition

Let $\boldsymbol{\theta}$ denote the parameters of the state space model. These parameters are embedded in the system matrices. For the state space model with a fixed value of $\boldsymbol{\theta}$, the Kalman Filter produces the prediction errors, $\mathbf{v}_t(\boldsymbol{\theta})$, and the prediction error variances, $\mathbf{F}_t(\boldsymbol{\theta})$, from the prediction equations. The *prediction error decomposition* of the log-likelihood function follows immediately:

$$\begin{aligned} \ln L(\boldsymbol{\theta}|\mathbf{y}) &= -\frac{NT}{2} \ln(2\pi) - \frac{1}{2} \sum_{t=1}^T \ln |\mathbf{F}_t(\boldsymbol{\theta})| \\ &\quad - \frac{1}{2} \sum_{t=1}^T \mathbf{v}_t'(\boldsymbol{\theta}) \mathbf{F}_t^{-1}(\boldsymbol{\theta}) \mathbf{v}_t(\boldsymbol{\theta}) \end{aligned}$$

Derivation of the Kalman Filter

The derivation of the Kalman filter equations relies on the following result:

Result 1: Suppose

$$\begin{pmatrix} \mathbf{y} \\ \mathbf{x} \end{pmatrix} \sim N \left(\begin{pmatrix} \boldsymbol{\mu}_y \\ \boldsymbol{\mu}_x \end{pmatrix}, \begin{pmatrix} \boldsymbol{\Sigma}_{yy} & \boldsymbol{\Sigma}_{yx} \\ \boldsymbol{\Sigma}_{xy} & \boldsymbol{\Sigma}_{xx} \end{pmatrix} \right)$$

Then, the distribution of \mathbf{y} given \mathbf{x} is observed is normal with

$$\begin{aligned} E[\mathbf{y}|\mathbf{x}] &= \boldsymbol{\mu}_{y|x} = \boldsymbol{\mu}_y + \boldsymbol{\Sigma}_{yx}\boldsymbol{\Sigma}_{xx}^{-1}(\mathbf{x} - \boldsymbol{\mu}_x) \\ \text{var}(\mathbf{y}|\mathbf{x}) &= \boldsymbol{\Sigma}_{yy} - \boldsymbol{\Sigma}_{yx}\boldsymbol{\Sigma}_{xx}^{-1}\boldsymbol{\Sigma}_{xy} \end{aligned}$$

Intuition: Think of the population regression of y on x in the univariate context

$$y = \alpha + \beta x + \varepsilon$$

Here

$$\alpha = \mu_y - \beta\mu_x, \quad \beta = \boldsymbol{\Sigma}_{yx}\boldsymbol{\Sigma}_{xx}^{-1}$$

Then

$$E[\mathbf{y}|\mathbf{x}] = \alpha + \beta x = \boldsymbol{\mu}_y + \boldsymbol{\Sigma}_{yx}\boldsymbol{\Sigma}_{xx}^{-1}(\mathbf{x} - \boldsymbol{\mu}_x)$$

In the linear Gaussian state-space model, the disturbances ε_t , and η_t are normally distributed and the initial state vector α_0 is also normally distributed. From the transition equation, the state vector at time 1 is

$$\alpha_1 = \mathbf{T}_1\alpha_0 + \mathbf{c}_1 + \mathbf{R}_1\eta_1$$

Since $\alpha_0 \sim N(\mathbf{a}_0, \mathbf{P}_0)$, $\eta_1 \sim N(\mathbf{0}, \mathbf{Q}_1)$ and α_0 and η_1 are independent it follows that

$$\begin{aligned} E[\alpha_1] &= \mathbf{a}_{1|0} = \mathbf{T}_1\mathbf{a}_0 + \mathbf{c}_1 \\ \text{var}(\alpha_1) &= E[(\alpha_1 - \mathbf{a}_{1|0})(\alpha_1 - \mathbf{a}_{1|0})'] \\ &= \mathbf{P}_{1|0} = \mathbf{T}_1\mathbf{P}_0\mathbf{T}_1' + \mathbf{R}_1\mathbf{Q}_1\mathbf{R}_1' \\ \alpha_1 &\sim N(\mathbf{a}_{1|0}, \mathbf{P}_{1|0}) \end{aligned}$$

Notice that the expression for $\mathbf{a}_{1|0}$ is the prediction equation for α_1 at $t = 0$.

Next, from the measurement equation

$$\mathbf{y}_1 = \mathbf{Z}_1\boldsymbol{\alpha}_1 + \mathbf{d}_1 + \boldsymbol{\varepsilon}_1$$

Since $\boldsymbol{\alpha}_1 \sim N(\mathbf{a}_{1|0}, \mathbf{P}_{1|0})$, $\boldsymbol{\varepsilon}_1 \sim N(\mathbf{0}, \mathbf{H}_1)$ and $\boldsymbol{\alpha}_1$ and $\boldsymbol{\varepsilon}_1$ are independent it follows that \mathbf{y}_1 is normally distributed with

$$\begin{aligned} E[\mathbf{y}_1] &= \mathbf{y}_{1|0} = \mathbf{Z}_1\mathbf{a}_{1|0} + \mathbf{d}_1 \\ \text{var}(\mathbf{y}_1) &= E[(\mathbf{y}_1 - \mathbf{y}_{1|0})(\mathbf{y}_1 - \mathbf{y}_{1|0})'] \\ &= E[(\mathbf{Z}_1(\boldsymbol{\alpha}_1 - \mathbf{a}_{1|0}) + \boldsymbol{\varepsilon}_1)(\mathbf{Z}_1(\boldsymbol{\alpha}_1 - \mathbf{a}_{1|0}) + \boldsymbol{\varepsilon}_1)'] \\ &= \mathbf{Z}_1\mathbf{P}_{1|0}\mathbf{Z}_1' + \mathbf{H}_1 \end{aligned}$$

Notice that the expression for $\mathbf{y}_{1|0}$ is the prediction equation for \mathbf{y}_1 at $t = 0$.

For the updating equations, the goal is to find the distribution of α_1 conditional on y_1 being observed. To do this, the joint normal distribution of (α_1', y_1') must be determined and then Result 1 can be applied. To determine the joint distribution of (α_1', y_1') use

$$\begin{aligned}\alpha_1 &= \mathbf{a}_{1|0} + (\alpha_1 - \mathbf{a}_{1|0}) \\ y_1 &= y_{1|0} + y_1 - y_{1|0} \\ &= \mathbf{Z}_1 \mathbf{a}_{1|0} + \mathbf{d}_1 + \mathbf{Z}_1 (\alpha_1 - \mathbf{a}_{1|0}) + \varepsilon_1\end{aligned}$$

and note that

$$\begin{aligned}\text{cov}(\alpha_1, y_1) &= E[(\alpha_1 - \mathbf{a}_{1|0})(y_1 - y_{1|0})'] \\ &= E[(\alpha_1 - \mathbf{a}_{1|0}) ((\mathbf{Z}_1 (\alpha_1 - \mathbf{a}_{1|0}) + \varepsilon_1)')] \\ &= E[(\alpha_1 - \mathbf{a}_{1|0}) ((\alpha_1 - \mathbf{a}_{1|0}) \mathbf{Z}_1' + \varepsilon_1')] \\ &= E[(\alpha_1 - \mathbf{a}_{1|0})(\alpha_1 - \mathbf{a}_{1|0}) \mathbf{Z}_1'] \\ &\quad + E[(\alpha_1 - \mathbf{a}_{1|0}) \varepsilon_1'] \\ &= \mathbf{P}_{1|0} \mathbf{Z}_1'\end{aligned}$$

Therefore,

$$\begin{pmatrix} \boldsymbol{\alpha}_1 \\ \mathbf{y}_1 \end{pmatrix} \sim N \left(\begin{pmatrix} \mathbf{a}_{1|0} \\ \mathbf{Z}_1 \mathbf{a}_{1|0} + \mathbf{d}_1 \end{pmatrix}, \begin{pmatrix} \mathbf{P}_{1|0} & \mathbf{P}_{1|0} \mathbf{Z}'_1 \\ \mathbf{Z}_1 \mathbf{P}_{1|0} & \mathbf{Z}_1 \mathbf{P}_{1|0} \mathbf{Z}'_1 + \mathbf{H}_1 \end{pmatrix} \right)$$

Now, use Result 1 to determine the mean and variance of the distribution of $\boldsymbol{\alpha}_1$ conditional on \mathbf{y}_1 being observed:

$$\begin{aligned} \boldsymbol{\alpha}_1 | \mathbf{y}_1 &\sim N(\mathbf{a}_1, \mathbf{P}_1) \\ \mathbf{a}_1 &= E[\boldsymbol{\alpha}_1 | \mathbf{y}_1] \\ &= \mathbf{a}_{1|0} + \mathbf{P}_{1|0} \mathbf{Z}'_1 (\mathbf{Z}_1 \mathbf{P}_{1|0} \mathbf{Z}'_1 + \mathbf{H}_1)^{-1} \\ &\quad \times (\mathbf{y}_1 - \mathbf{Z}_1 \mathbf{a}_{1|0} - \mathbf{d}_1) \\ &= \mathbf{a}_{1|0} + \mathbf{P}_{1|0} \mathbf{Z}'_1 \mathbf{F}_1^{-1} \mathbf{v}_1 \\ \mathbf{P}_1 &= \text{var}(\boldsymbol{\alpha}_1 | \mathbf{y}_1) \\ &= \mathbf{P}_{1|0} - \mathbf{P}_{1|0} \mathbf{Z}'_1 (\mathbf{Z}_1 \mathbf{P}_{1|0} \mathbf{Z}'_1 + \mathbf{H}_1)^{-1} \mathbf{Z}_1 \mathbf{P}_{1|0} \\ &= \mathbf{P}_{1|0} - \mathbf{P}_{1|0} \mathbf{Z}'_1 \mathbf{F}_1^{-1} \mathbf{Z}_1 \mathbf{P}_{1|0} \end{aligned}$$

where

$$\begin{aligned}\mathbf{v}_1 &= \mathbf{y}_1 - \mathbf{y}_{1|0} = (\mathbf{y}_1 - \mathbf{Z}_1 \mathbf{a}_{1|0} - \mathbf{d}_1) \\ \mathbf{F}_1 &= E[\mathbf{v}_1 \mathbf{v}'_1] = \mathbf{Z}_1 \mathbf{P}_{1|0} \mathbf{Z}'_1 + \mathbf{H}_1\end{aligned}$$

Notice that the expressions for \mathbf{a}_1 and \mathbf{P}_1 are exactly the Kalman filter updating equations for $t = 1$. Repeating the above prediction and updating steps for $t = 2, \dots, T$ gives the Kalman filter recursion equations.

Recursive least Squares Estimation and Tests for Structural Stability

Consider the CAPM regression

$$\begin{aligned}y_t &= \alpha + \beta_M r_{M,t} + \xi_t, \quad \xi_t \sim N(0, \sigma_\xi^2) \\ &= \mathbf{x}'_t \boldsymbol{\beta} + \xi_t \\ \mathbf{x}_t &= (1, r_{M,t})', \quad \boldsymbol{\beta} = (\alpha, \beta_M)'\end{aligned}$$

Define $\boldsymbol{\alpha}_t = \boldsymbol{\alpha}_{t-1} = \boldsymbol{\beta}$. Then the transition equation is

$$\boldsymbol{\alpha}_t = \mathbf{I}_2 \boldsymbol{\alpha}_{t-1}$$

and the measurement equation is

$$y_t = \mathbf{x}'_t \boldsymbol{\alpha}_t + \sigma_\xi \varepsilon_t$$

That is,

$$\mathbf{T}_t = \mathbf{I}_2, \mathbf{Z}_t = \mathbf{x}'_t, \mathbf{H}_t = \sigma_\xi, \mathbf{R}_t = 0.$$

The coefficient vector $\boldsymbol{\beta}$ is fixed and unknown so that the initial conditions are $\boldsymbol{\alpha}_1 \sim N(\mathbf{0}, \kappa \mathbf{I}_k)$ where κ is large.

Recursive least squares estimation

The *recursive least squares* (RLS) estimates of the regression coefficient vector β are readily computed from the Kalman filter. The RLS estimates are based on estimating the model

$$y_t = \beta'_t \mathbf{x}_t + \xi_t, \quad t = 1, \dots, n$$

by least squares recursively for $t = 3, \dots, n$ giving $n - 2$ least squares (RLS) estimates $(\hat{\beta}_3, \dots, \hat{\beta}_T)$.

1. If β is constant over time then the recursive estimates $\hat{\beta}_t$ should quickly settle down near a common value.
2. If some of the elements in β are not constant then the corresponding RLS estimates should show instability. Hence, a simple graphical technique for uncovering parameter instability is to plot the RLS estimates $\hat{\beta}_{it}$ ($i = 1, 2$) and look for instability in the plots.

Tests for constant parameters

Formal tests for structural stability of the regression coefficients, such as the CUSUM test of Brown, Durbin and Evans (1976), may be computed from the standardized 1 – *step* ahead recursive residuals

$$w_t = \frac{v_t}{\sqrt{f_t}} = \frac{y_t - \hat{\beta}'_{t-1} \mathbf{x}_t}{\sqrt{f_t}}$$

These standardized recursive residuals result as a by-product of the Kalman filter recursions

The CUSUM test is based on

$$\text{CUSUM}_t = \sum_{j=k+1}^t \frac{\hat{w}_j}{\hat{\sigma}_w}$$

where $\hat{\sigma}_w$ is the sample standard deviation of \hat{w}_j and k denotes the number of estimated coefficients. Under the null hypothesis that β is constant, CUSUM_t has mean zero and variance that is proportional to $t - k - 1$.