# Financial Econometrics
# Kalman Filter: some applications to Finance
# University of Evry - Master 2

Eric Bouyé*

January 27, 2009

## Contents



*University of Evry and Société Générale. Contact: eric.bouye@sgam.com

# 1 State-space models

A general specification for the Kalman filter - KALMAN (1960), [6]) - is provided below. The representation of HARVEY (1991) [5] is used. For an excellent presentation about filtering - and elegant proofs-, we refer to LE GLAND (2007).

$$\begin{cases} \text{Measurement equation} & y_t = \mathbf{X}_t \beta_t + d_t + \varepsilon_t & t = 1, \ldots, T \\ \\ \text{Transition equation} & \beta_t = \mathbf{T}_t \beta_{t-1} + c_t + \mathbf{R}_t \eta_t & t = 1, \ldots, T \end{cases} \tag{1}$$

with $\beta_t$ the $(m \times 1)$ state vector (unobserved) that represents the development over time of the system. $y_t$ is the $(N \times 1)$ vector of the observed variables. $d_t$ and $\varepsilon_t$ are $(N \times 1)$ vectors and $\mathbf{X}_t$ is a $(N \times m)$ matrix. $\mathbf{T}_t$ is an $(m \times m)$ matrix, $\mathbf{R}_t$ a $(m \times g)$ matrix, $\mathbf{c}_t$ a $(m \times 1)$ vector and $\boldsymbol{\eta}_t$ a $(g \times 1)$ vector. The disturbances are such that

$$\begin{cases} \varepsilon_t \sim \mathcal{N}(0, \mathbf{H}_t) \\ \eta_t \sim \mathcal{N}(0, \mathbf{Q}_t) \end{cases}$$

and $\varepsilon_t$ and $\eta_t$ are uncorrelated. Specifically,

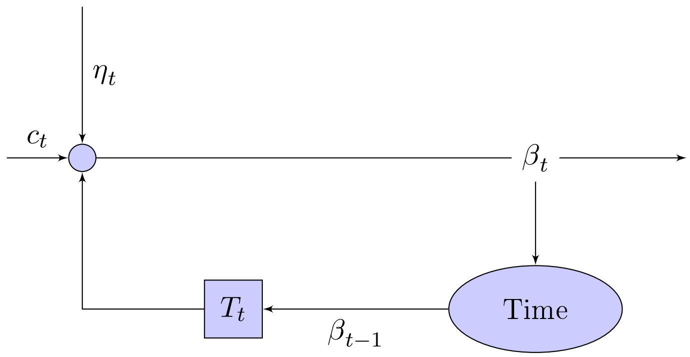
$$\begin{cases} \mathbb{E}(\varepsilon_t \eta_s^\intercal) = 0 & \forall s, t = 1, \ldots, T \\ \mathbb{E}(\varepsilon_t \beta_0^\intercal) = 0 & \text{for } t = 1, \ldots, T \end{cases} .$$

# 2 The Scalar Kalman Filter

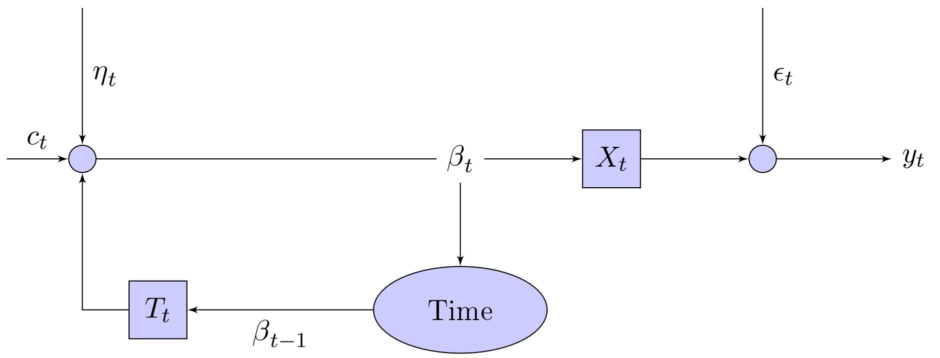
## 2.1 Presentation

For the sake of simplicity, the scalar model is first studied

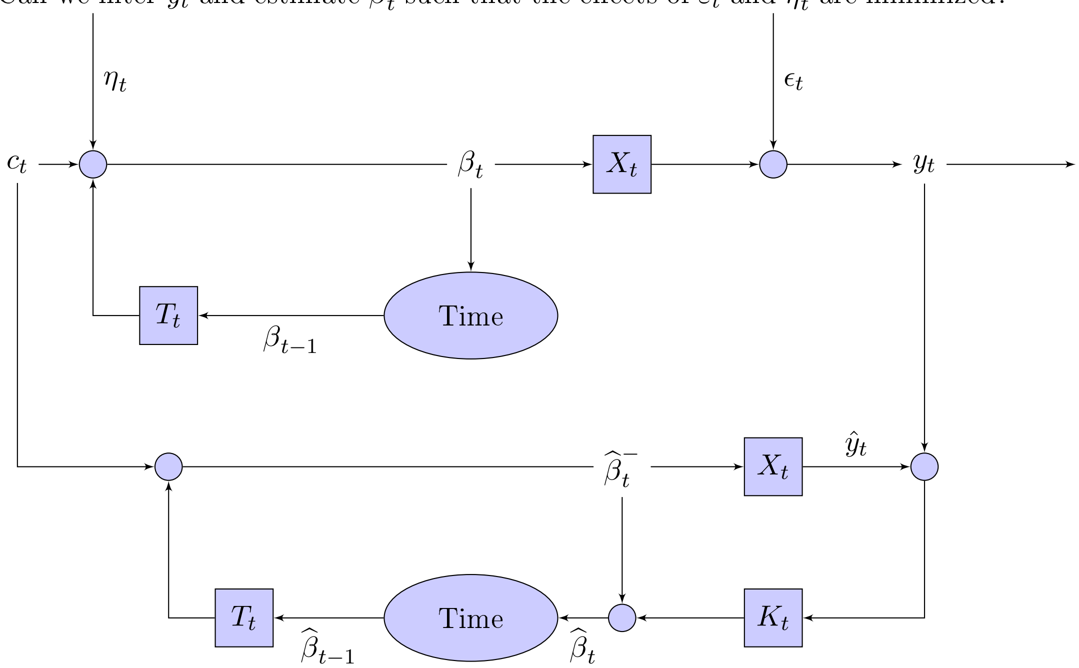
$$\begin{cases} \text{Measurement equation} & y_t = X_t \beta_t + \varepsilon_t & t = 1, \ldots, T \\ \\ \text{Transition equation} & \beta_t = T_t \beta_{t-1} + c_t + \eta_t & t = 1, \ldots, T \end{cases}$$




Can we filter $\beta_t$ such that the effects of $\eta_t$ are minimized?




Can we filter $y_t$ and estimate $\beta_t$ such that the effects of $\varepsilon_t$ and $\eta_t$ are minimized?




In order to predict an estimate of the output $y_t$, the KF uses an a priori estimate of $\beta_t$ noted $\widehat{\beta}_t^-$ such that

$$\widehat{\beta}_t^- = T_t \widehat{\beta}_{t-1} + c_t.$$

Then the difference between the estimated output $\widehat{y}_t$ and actual output can be computed

$$y - \widehat{y}_t = y_t - X_t \widehat{\beta}_t^-$$

and this residual is used to refine the a priori estimate: the a posteriori estimate $\widehat{\beta}_t$ is calculated

$$\begin{aligned} \widehat{\beta}_t &= \widehat{\beta}_t^- + K_t \left( y - \widehat{y}_t \right) \\ &= \widehat{\beta}_t^- + K_t \left( y_t - X_t \widehat{\beta}_t^- \right). \end{aligned}$$

**Question: what is the appropriate $K_t$ such that an optimal estimator is computed ?**

The errors are defined

$$\begin{cases} e_t^- = \beta_t - \widehat{\beta}_t^- & \text{a priori} \\ e_t = \beta_t - \widehat{\beta}_t & \text{a posteriori} \end{cases}$$

with respective variances

$$\begin{cases} p_t^- = E\left(\left(e_t^-\right)^2\right) \\ p_t = E\left(e_t^2\right). \end{cases}$$

**A Kalman Filter minimizes the a posteriori error variance $p_t$**

Then, let us compute the value of $K_t$ that minimizes $p_t$ i.e.

$$\begin{aligned} K^\star &= \arg\min p_t(K) \\ &= \arg\min \mathbf{E}\left[\left(\beta_t - \widehat{\beta}_t\right)^2\right] \\ &= \arg\min \mathbf{E}\left[\left(\beta_t - \widehat{\beta}_t^- - K\left(y_t - X_t\widehat{\beta}_t^-\right)\right)^2\right] \end{aligned}$$

Then

$$\frac{\partial p_t(K)}{\partial K} = 0$$
$$\Updownarrow$$
$$\cdots$$
$$\Updownarrow$$
$$K^\star = \frac{X_t p_t^-}{X_t^2 p_t^- + H_t}.$$

From this result, one can remark that

- if the a priori error is large, then

$$K^\star \longrightarrow \frac{1}{X_t}$$

  and the confidence in the a priori estimate is quite low,

- if the a priori error is small, then the correction by the KF will be small. In other words, if our first estimates have been quite good, there is little need to correct it,

- if $H_t$ is large, the confidence in the measurement is quite low. The KF has more confidence in the previous estimates.

### 2.2 Summary of the different steps

1. **Model specification**

$$\begin{cases} \text{Measurement equation} & y_t = X_t\beta_t + \varepsilon_t & t = 1, \ldots, T \\ \text{Transition equation} & \beta_t = T_t\beta_{t-1} + c_t + \eta_t & t = 1, \ldots, T \end{cases}$$

2. **Prediction step**

$$\widehat{\beta}_t^- = T_t\widehat{\beta}_{t-1} + c_t.$$

and the a priori covariance is

$$p_t^- = T_t^2 p_{t-1} + Q$$

3. **Correction step**

Given the additional information on the output, one can correct the estimator

$$\widehat{\beta}_t = \widehat{\beta}_t^- + K_t\left(y_t - X_t\widehat{\beta}_t^-\right)$$

with $K_t$ the gain:

$$K_t = \frac{X_t p_t^-}{X_t^2 p_t^- + H_t}$$

and the a posteriori covariance is

$$p_t = p_t^-\left(1 - X_t K_t\right).$$

### 2.3 Is it an optimal linear estimate ?

Is the Kalman Filter an optimal **linear** estimate ?

$$\hat{\beta}_t = a_1\hat{\beta}_{t-1} + a_2 y_t$$

$$\min p_t$$
$$\min \mathbf{E}\left[\left(\beta_t - \widehat{\beta}_t\right)^2\right]$$

$$\begin{cases} \frac{\partial p_t}{\partial a_1} = 0 \\ \frac{\partial p_t}{\partial a_2} = 0 \end{cases} \Longleftrightarrow \cdots \Longleftrightarrow \begin{cases} \mathbf{E}\left[\left(\beta_t - \widehat{\beta}_t\right)\widehat{\beta}_{t-1}\right] = 0 \\ \mathbf{E}\left[\left(\beta_t - \widehat{\beta}_t\right)y_t\right] = 0 \end{cases}$$

that corresponds to the orthogonality conditions discussed above. Then we obtain

$$a_1 = T_t\left(1 - a_2 X_t\right)$$

Then, by substituting in the above equation

$$\begin{aligned}\hat{\beta}_t &= a_1\hat{\beta}_{t-1} + a_2 y_t \\ &= T_t\left(1 - a_2 X_t\right)\hat{\beta}_{t-1} + a_2 y_t \\ &= T_t\hat{\beta}_{t-1} + a_2\left(y_t - T_t X_t \hat{\beta}_{t-1}\right) \\ \hat{\beta}_t &= \hat{\beta}_t^- + a_2\left(y_t - X_t\widehat{\beta}_t^-\right).\end{aligned}$$

### 2.4 Prediction, filtering and smoothing

One distinguishes three types of use of the Kalman filter

1. Prediction
2. Filtering
3. Smoothing

## 3 The Kalman Filter

For the sake of simplicity, let denote $\mathbb{E}\left(x_t^-\right) = \mathbb{E}\left(x_t \mid y_0, \ldots, y_{t-1}\right)$ and $\mathbb{E}\left(x_t\right) = \mathbb{E}\left(x_t \mid y_0, \ldots, y_t\right)$.

The state-space model is:

$$\begin{cases} \text{Measurement equation} & y_t = \mathbf{X}_t\beta_t + d_t + \varepsilon_t & t = 1, \ldots, T \\ \text{Transition equation} & \beta_t = \mathbf{T}_t\beta_{t-1} + c_t + \mathbf{R}_t\eta_t & t = 1, \ldots, T \end{cases}$$

Theorem *The Kalman-Buci filter.*

The Kalman-Buci filter is defined as follows

$$\mathbb{E}\left(\beta_t^-\right) = \mathbf{T}_t\mathbb{E}\left(\beta_{t-1}\right) + c_t \tag{2}$$

$$\mathbf{P}_t^- = \mathbf{T}_t\mathbf{P}_{t-1}\mathbf{T}_t^\intercal + \mathbf{R}_t\mathbf{Q}_t\mathbf{R}_t^\intercal \tag{3}$$

and

$$\mathbb{E}\left(\beta_t\right) = \mathbb{E}\left(\beta_t^-\right) + \mathbf{K}_t\left[y_t - \left(\mathbf{X}_t\mathbb{E}\left(\beta_t^-\right) + d_t\right)\right] \tag{4}$$

$$\mathbf{P}_t = \left[\mathbf{I} - \mathbf{K}_t\mathbf{X}_t\right]\mathbf{P}_t^- \tag{5}$$

with $\mathbf{K}_t$ the Kalman gain:

$$\mathbf{K}_t = \mathbf{P}_t^-\mathbf{X}_t^\intercal\left[\mathbf{X}_t\mathbf{P}_t^-\mathbf{X}_t^\intercal + \mathbf{Q}_t\right]^{-1} \tag{6}$$

The initial state of the system is as follows

$$\left\{\begin{array}{l}\mathbb{E}\left(\beta_0^-\right)=\mathbb{E}\left(\beta_0\right)\\ \mathbf{P}_0^-=\mathbf{Var}\left(\beta_0\right)=\mathbf{P}_0\end{array}\right. .$$

Proof for (2).

$$\begin{aligned}\mathbb{E}\left(\beta_t^-\right) &= \mathbf{T}_t\mathbb{E}\left(\beta_{t-1}^-\right)+c_t+\mathbf{R}_t\mathbb{E}\left(\eta_t^-\right)\\ \mathbb{E}\left(\beta_t^-\right) &= \mathbf{T}_t\mathbb{E}\left(\beta_{t-1}\right)+c_t\end{aligned}$$

Proof for (3).

$$\begin{aligned}\mathbf{P}_t^- &= \mathbb{E}\left[\left(\beta_t-\mathbb{E}\left(\beta_t^-\right)\right)\left(\beta_t-\mathbb{E}\left(\beta_t^-\right)\right)^\intercal\right]\\ &= \mathbb{E}\left\{\left[\mathbf{T}_t\left(\beta_{t-1}-\mathbb{E}\left(\beta_{t-1}\right)\right)+\mathbf{R}_t\mathbb{E}\left(\eta_t\right)\right]\left[\mathbf{T}_t\left(\beta_{t-1}-\mathbb{E}\left(\beta_{t-1}\right)\right)+\mathbf{R}_t\mathbb{E}\left(\eta_t\right)\right]^\intercal\right\}\\ &= \mathbf{T}_t\mathbb{E}\left\{\left(\beta_{t-1}-\mathbb{E}\left(\beta_{t-1}\right)\right)\left(\beta_{t-1}-\mathbb{E}\left(\beta_{t-1}\right)\right)^\intercal\right\}\mathbf{T}_t^\intercal+\mathbf{R}_t\mathbb{E}\left(\eta_t\eta_t^\intercal\right)\mathbf{R}_t^\intercal\\ &\quad+\mathbf{T}_t\mathbb{E}\left\{\left(\beta_{t-1}-\mathbb{E}\left(\beta_{t-1}\right)\right)\eta_t^\intercal\right\}\mathbf{T}_t^\intercal+\mathbf{R}_t\mathbb{E}\left\{\eta_t\left(\beta_{t-1}-\mathbb{E}\left(\beta_{t-1}\right)\right)^\intercal\right\}\mathbf{T}_t^\intercal\end{aligned}$$

By noting that $\mathbb{E}\left\{\left(\beta_{t-1}-\mathbb{E}\left(\beta_{t-1}\right)\right)\eta_t^\intercal\right\}=0$, we get

$$\mathbf{P}_t^-=\mathbf{T}_t\mathbf{P}_{t-1}\mathbf{T}_t^\intercal+\mathbf{R}_t\mathbf{Q}_t\mathbf{R}_t^\intercal.$$

Proof for (4) and (5).
Let us note $\nu_t$ the innovation vector

$$\nu_t=y_t-\mathbb{E}\left(y_t\mid y_0,\ldots,y_{t-1}\right)=y_t-E\left(y_t^-\right).$$

We have

$$\begin{aligned}\mathbb{E}\left(\beta_t\mid y_0,\ldots,y_t\right) &= \mathbb{E}\left(\beta_t^-\right)+\mathbb{E}\left(\beta_t-\beta_t^-\mid y_0,\ldots,y_t\right)\\ &= \mathbb{E}\left(\beta_t^-\right)+\mathbb{E}\left(\beta_t-\beta_t^-\mid y_0,\ldots,y_{t-1},\nu_t\right)\\ \mathbb{E}\left(\beta_t\right) &= \mathbb{E}\left(\beta_t^-\right)+\mathbb{E}\left(\beta_t-\beta_t^-\mid\nu_t\right).\end{aligned}$$

From above,

$$\begin{aligned}\beta_t-\mathbb{E}\left(\beta_t\right) &= v-\left(\mathbb{E}\left(\beta_t\right)-\mathbb{E}\left(\beta_t^-\right)\right)\\ &= \left(\beta_t-\mathbb{E}\left(\beta_t^-\right)\right)-\mathbb{E}\left(\beta_t-\beta_t^-\mid\nu_t\right)\end{aligned}$$

then, we can calculate $\mathbf{P}_t$:

$$\begin{aligned}\mathbf{P}_t &= \mathbb{E}\left[\left(\beta_t-\mathbb{E}\left(\beta_t\right)\right)\left(\beta_t-\mathbb{E}\left(\beta_t\right)\right)^\intercal\right]\\ &= \mathbb{E}\left[\left(\beta_t-\mathbb{E}\left(\beta_t^-\right)\right)-\mathbb{E}\left(\beta_t-\beta_t^-\mid\nu_t\right)\left(\beta_t-\mathbb{E}\left(\beta_t^-\right)\right)-\mathbb{E}\left(\beta_t-\beta_t^-\mid\nu_t\right)^\intercal\right]\end{aligned}$$

We need to compute the conditial mean and variance of $\left(\beta_t - \mathbb{E}\left(\beta_t^-\right)\right)$. First, we need to compute the variance of $\nu_t$ and its covariance with $\left(\beta_t - \mathbb{E}\left(\beta_t^-\right)\right)$. We have

$$
\begin{aligned}
\mathbb{E}\left(\nu_t \nu_t^{\intercal}\right) &= \mathbb{E}\left[\left(y_t - E\left(y_t^-\right)\right)\left(y_t - E\left(y_t^-\right)\right)^{\intercal}\right] \\
&= \mathbb{E}\left[\left(\mathbf{X}_t\left(\beta_t - \mathbb{E}\left(\beta_t^-\right)\right) + \varepsilon_t\right)\left(\mathbf{X}_t\left(\beta_t - \mathbb{E}\left(\beta_t^-\right)\right) + \varepsilon_t\right)^{\intercal}\right] \\
&= \mathbf{X}_t \mathbb{E}\left[\left(\beta_t - \mathbb{E}\left(\beta_t^-\right)\right)\left(\beta_t - \mathbb{E}\left(\beta_t^-\right)\right)^{\intercal}\right]\mathbf{X}_t^{\intercal} + \mathbb{E}\left[\varepsilon_t\left(\beta_t - \mathbb{E}\left(\beta_t^-\right)\right)\right]\mathbf{X}_t^{\intercal} \\
&\quad + \mathbf{X}_t \mathbb{E}\left[\left(\beta_t - \mathbb{E}\left(\beta_t^-\right)\right)\varepsilon_t^{\intercal}\right] + \mathbb{E}\left(\varepsilon_t \varepsilon_t^{\intercal}\right) \\
\mathbb{E}\left(\nu_t \nu_t^{\intercal}\right) &= \mathbf{X}_t \mathbf{P}_t^- \mathbf{X}_t^{\intercal} + \mathbf{H}_t
\end{aligned}
$$

Also the covariance is such as

$$
\begin{aligned}
\mathbb{E}\left(\left(\beta_t - \mathbb{E}\left(\beta_t^-\right)\right)\nu_t^{\intercal}\right) &= \mathbb{E}\left(\left(\beta_t - \mathbb{E}\left(\beta_t^-\right)\right)\left(\mathbf{X}_t\left(\beta_t - \mathbb{E}\left(\beta_t^-\right)\right) + \varepsilon_t\right)^{\intercal}\right) \\
&= \mathbb{E}\left[\left(\beta_t - \mathbb{E}\left(\beta_t^-\right)\right)\left(\beta_t - \mathbb{E}\left(\beta_t^-\right)\right)^{\intercal}\right]\mathbf{X}_t^{\intercal} + \mathbb{E}\left[\left(\beta_t - \mathbb{E}\left(\beta_t^-\right)\right)\varepsilon_t^{\intercal}\right] \\
\mathbb{E}\left(\nu_t \nu_t^{\intercal}\right) &= \mathbf{P}_t^- \mathbf{X}_t^{\intercal}.
\end{aligned}
$$

Then, the joint distribution of $\left(\beta_t - \mathbb{E}\left(\beta_t\right)\right)$ and $\nu_t$ is

$$
\begin{pmatrix} \beta_t - \mathbb{E}\left(\beta_t\right) \\ \nu_t \end{pmatrix} \sim \mathcal{N}\left(\mathbf{0}, \begin{pmatrix} \mathbf{P}_t^- & \mathbf{P}_t^- \mathbf{X}_t^{\intercal} \\ \mathbf{X}_t \mathbf{P}_t^- & \mathbf{X}_t \mathbf{P}_t^- \mathbf{X}_t^{\intercal} + \mathbf{H}_t \end{pmatrix}\right)
$$

then if $\mathbf{Q}_t$ is invertible, then $\mathbf{X}_t \mathbf{P}_t^- \mathbf{X}_t^{\intercal} + \mathbf{H}_t$ is also invertible, and from a property of the gaussian conditional distribution (see Appendix 2),

$$
\mathbb{E}\left(\beta_t\right) = \mathbb{E}\left(\beta_t^-\right) + \mathbf{P}_t^- \mathbf{X}_t^{\intercal}\left[\mathbf{X}_t \mathbf{P}_t^- \mathbf{X}_t^{\intercal} + \mathbf{H}_t\right]^{-1}\nu_t
$$

and

$$
\mathbf{P}_t = \mathbf{P}_t^- - \mathbf{P}_{\mathbf{t}}^- \mathbf{X}_{\mathbf{t}}^{\intercal}\left[\mathbf{X}_{\mathbf{t}} \mathbf{P}_{\mathbf{t}}^- \mathbf{X}_{\mathbf{t}}^{\intercal} + \mathbf{H}_{\mathbf{t}}\right]^{-1}\mathbf{X}_t \mathbf{P}_{\mathbf{t}}^-.
$$

**Set of equations for the Kalman Filter: alternative notation**

$$
\left\{
\begin{array}{l}
\mathbb{E}\left(\beta_t^-\right) = \mathbf{T}_t \mathbb{E}\left(\beta_{t-1}\right) + c_t \\
\mathbf{P}_t^- = \mathbf{T}_t \mathbf{P}_{t-1} \mathbf{T}_t^{\intercal} + \mathbf{R}_t \mathbf{Q}_t \mathbf{R}_t^{\intercal} \\
E\left(y_t^-\right) = \mathbf{X}_t \mathbb{E}\left(\beta_t^-\right) + d_t \\
\nu_t = y_t - E\left(y_t^-\right) \\
\mathbb{E}\left(V_t\right) = \mathbf{X}_t \mathbf{P}_t^- \mathbf{X}_t^{\intercal} + \mathbf{Q}_t \\
\mathbb{E}\left(\beta_t\right) = \mathbb{E}\left(\beta_t^-\right) + \mathbf{P}_t^- \mathbf{X}_t^{\intercal}\left[\mathbb{E}\left(V_t\right)\right]^{-1}\nu_t \\
\mathbf{P}_t = \mathbf{P}_t^- - \mathbf{P}_{\mathbf{t}}^- \mathbf{X}_{\mathbf{t}}^{\intercal}\left[\mathbb{E}\left(\mathbf{V}_{\mathbf{t}}\right)\right]^{-1}\mathbf{X}_t \mathbf{P}_t^-
\end{array}
\right.
\iff
\left\{
\begin{array}{l}
\hat{\beta}_{t|t-1} = \mathbf{T}_t \hat{\beta}_{t-1|t-1} + c_t \\
\hat{\mathbf{P}}_{t|t-1} = \mathbf{T}_t \hat{\mathbf{P}}_{t-1|t-1} \mathbf{T}_t^{\intercal} + \mathbf{R}_t \mathbf{Q}_t \mathbf{R}_t^{\intercal} \\
\hat{y}_{t|t-1} = \mathbf{X}_t \hat{\beta}_{t|t-1} + d_t \\
\hat{\boldsymbol{\nu}}_t = y_t - \hat{y}_{t|t-1} \\
\hat{V}_t = \mathbf{X}_t \hat{\mathbf{P}}_{t|t-1} \mathbf{X}_t^{\intercal} + \mathbf{Q}_t \\
\hat{\beta}_{t|t} = \hat{\beta}_{t|t-1} + \hat{\mathbf{P}}_{t|t-1} \mathbf{X}_t^{\intercal} \hat{V}_t^{-1} \boldsymbol{\nu}_t \\
\mathbf{P}_{t|t} = \hat{\mathbf{P}}_{t|t-1} - \hat{\mathbf{P}}_{\mathbf{t|t-1}} \mathbf{X}_t \hat{\mathbf{V}}_{\mathbf{t}}^{-1} \mathbf{X}_t \hat{\mathbf{P}}_{t|t-1}
\end{array}
\right.
$$

# 4 Generalized Vasicek Term Structure Models

## 4.1 The Model

From Babbs and Ben Nowman (1999). Their model is a subclass of Langetieg's linear Gaussian models of the term structure:

$$B(M,t)=\exp\left\{-\tau\left[R(\infty)-w(\tau)-\sum_{j=1}^{J}u(\xi_j\tau)X_j(\tau)\right]\right\}$$

with

$$R(\infty)=\mu+\sum_{q=1}^{Q}\theta_q\sum_{j=1}^{J}\frac{\kappa_{jq}}{\xi_j}-\frac{1}{2}\sum_{q=1}^{Q}\left(\sum_{j=1}^{J}\frac{\kappa_{jq}}{\xi_j}\right)^2$$

$$\begin{aligned}W(\tau) &= \sum_{j=1}^{J}u(\xi_j\tau)\left[\sum_{q=1}^{Q}\theta_q\frac{\kappa_{jq}}{\xi_j}-\sum_{q=1}^{Q}\sum_{i=1}^{J}\frac{\kappa_{iq}\kappa_{jq}}{\xi_i\xi_j}\right]\\ &\quad+\frac{1}{2}\sum_{i=1}^{J}\sum_{j=1}^{J}H((\xi_i+\xi_j)\tau)\sum_{q=1}^{Q}\frac{\kappa_{iq}\kappa_{jq}}{\xi_i\xi_j}\end{aligned}$$

with

$$\tau=M-t\qquad\text{and}\qquad u(x)=\frac{1-e^{-x}}{x}.$$

The theoretical yield curve is such that

$$\begin{aligned}R(t+\tau_i,t) &= -\frac{\log B(t+\tau_i,t)}{\tau}\\ &= A_0(\tau_i)-A_1(\tau_i)^\top X(t)\qquad\text{for } i=1,\ldots,N\end{aligned}$$

with $A_0(\tau_i)=R(\infty)-w(\tau_i)$ and $A_1(\tau_i)=u(\xi_j\tau_i)$ a $J\times 1$ vector.

The time invariant state-space model is specified as follows:

$$\begin{cases}\text{Measurement equation} & R_t=\mathbf{Z}(\psi)X_t+d(\psi)+\varepsilon_t \qquad t=1,\ldots,T\\ \text{Transition equation} & X_t=\mathbf{T}(\psi)X_{t-1}+\eta_t \qquad t=1,\ldots,T\end{cases}$$

with $\varepsilon_t\sim\mathcal{N}(0,H(\psi))$ and $\eta_t\sim\mathcal{N}(0,V(\psi))$,

$$d(\psi)=\begin{pmatrix}A_0(\tau_1,\psi)\\ A_0(\tau_2,\psi)\\ \vdots\\ A_0(\tau_N,\psi)\end{pmatrix},$$

$$
\mathbf{Z}(\psi) = \begin{pmatrix} -A_1(\tau_1,\psi)^\top \\ -A_1(\tau_2,\psi)^\top \\ \vdots \\ -A_1(\tau_N,\psi)^\top \end{pmatrix} = \begin{pmatrix} -u(\xi_1\tau_1) & -u(\xi_2\tau_1) & \cdots & -u(\xi_J\tau_1) \\ -u(\xi_1\tau_2) & -u(\xi_2\tau_2) & \cdots & -u(\xi_J\tau_2) \\ \vdots & \vdots & \ddots & \vdots \\ -u(\xi_1\tau_N) & -u(\xi_2\tau_N) & \cdots & -u(\xi_J\tau_N) \end{pmatrix}.
$$

and

$$
\mathbf{T}(\psi) = \begin{pmatrix} e^{-\xi_1\Delta t} \\ e^{-\xi_2\Delta t} \\ \vdots \\ e^{-\xi_J\Delta t} \end{pmatrix},
$$

In practice, $H(\psi)$ is assumed to be diagonal with specific variances

$$
H(\psi) = \begin{pmatrix} h_1 & 0 & \cdots & 0 \\ 0 & h_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & h_N \end{pmatrix}.
$$

### 4.2 Empirical application

## 5 Investment Strategies Tracking problems

For this application we refer to Roncalli and Weisang (2008). Let assume we want to replicate the performance of a fund or an index $r_t^F$ (for example not investable) with some liquid instruments (e.g. futures) that are called factors. Let assume that we have $m$ factors with respective returns $r_t^i$ for $i = 1, \ldots, m$. It is assumed that the performance of the fund can be linearly explained by the returns of the factor. One can define the following tracking problem:

$$
\begin{cases} r_t^F = \mathbf{r}_t^\intercal \mathbf{w}_t + \eta_t \\ \mathbf{w}_t = \mathbf{w}_{t-1} + \varepsilon_t \end{cases}. \tag{7}
$$

In the above TP, the prediction part of the KF algorithm is used in practice.

### 5.1 Innovation representation

Let us define the **innovation representation** of the state-space model. We note that

$$
\begin{aligned}
\hat{\beta}_{t+1|t} &= (\mathbf{T}_{t+1} - \mathbf{K}_t\mathbf{X}_t)\,\hat{\beta}_{t|t-1} + \mathbf{K}_t y_t + (c_{t+1} - K_t d_t) \\
&= \mathbf{T}_{t+1}\hat{\beta}_{t|t-1} + c_{t+1} + \mathbf{K}_t\left(y_t - \mathbf{X}_t\hat{\beta}_{t|t-1} - d_t\right)
\end{aligned}
$$

with $\mathbf{K}_t = \mathbf{T}_{t+1}\hat{\mathbf{P}}_{\mathbf{t|t-1}}\mathbf{X_t^\intercal}\hat{\mathbf{V}}_{\mathbf{t}}^{\mathbf{-1}}$. Then,

$$\begin{cases} y_t = d_t + \mathbf{X}_t\hat{\beta}_{t|t-1} + \boldsymbol{\nu}_t & t = 1, \ldots, T \\ \\ \hat{\beta}_{t+1|t} = c_{t+1} + \mathbf{T}_{t+1}\hat{\beta}_{t|t-1} + \mathbf{K}_t\boldsymbol{\nu}_t & t = 1, \ldots, T \end{cases}.$$

Appliying this representation to our TP problem above, we get

$$\begin{cases} r_t^F = \mathbf{r}_t^\intercal \hat{\mathbf{w}}_{t|t-1} + \hat{\nu}_t \\ \hat{\mathbf{w}}_{t+1|t} = \hat{\mathbf{w}}_{t|t-1} + \hat{\mathbf{P}}_{\mathbf{t|t-1}}\mathbf{r}_t \left(\frac{\hat{\nu}_t}{\hat{V}_t}\right) \end{cases}. \tag{8}$$

with $\left(\frac{\hat{\nu}_t}{\hat{V}_t}\right)$ the normalized tracking-error. The $i^{th}$ factor is then ajusted as follows

$$\Delta\hat{w}^i_{t+1|t} = \hat{w}^i_{t+1|t} - \hat{w}^i_{t|t-1} \tag{9}$$

$$\Delta\hat{w}^i_{t+1|t} = \left(\frac{\hat{\nu}_t}{\hat{V}_t}\right)\sum_{j=1}^{m}\left(\hat{\mathbf{P}}_{\mathbf{t|t-1}}\right)_{i,j} r_k^j. \tag{10}$$

**Comments**.

### 5.2 Empirical application

## References


[1] Babbs Simon H. and K. Ben Nowman (1999), Kalman Filtering of Generalized Vasicek Term Structure Models, Journal of Financial and Quantitative Analysis, **34**, No. 1, March.

[2] Cheever Erik, Kalman Filter Tutorial, Swarthmore College.

[3] Davidson Russell and James G. MacKinnon (1993), *Estimation and Inference in Econometrics*, Oxford University Press (New York).

[4] Davidson Russell and James G. MacKinnon (2003), *Econometric Theory and Methods*, Oxford University Press (New York).

[5] Harvey A. C. (1991), *Forecasting, Structural Time Series and the Kalman Filter*, Cambridge: Cambridge University.

[6] Kalman R.E. (1960), A new approach to linear filtering and prediction problems, *J. Basic Eng.*, **82**, 35-45, March.

[7] Le Gland François (2007), Filtre de Kalman et modèles de Markov cachés, *Notes de cours.*

[8] Roncalli Thierry (2005), Econométrie Financière II, *Lecture Notes 2005-2006*, Master Ingénierie Economique et Statistique Appliquée, Université d'Evry.

[9] Roncalli Thierry and Guillaume Weisang (2008), Tracking Problems, Hedge Fund Replication and Alternative Beta, *Working Paper.*

[10] Welch Greg and Gary Bishop (2006), An Introduction to the Kalman Filter, *Working Paper.*

# 6 Appendix: The conditional gaussian distribution

Let assume a gaussian vector $(X_1, X_2)$ such as

$$(X_1, X_2) \sim \mathcal{N}\left(\left(\begin{array}{c} \mu_1 \\ \mu_2 \end{array}\right), \left(\begin{array}{cc} \Sigma_1 & \Sigma_{12} \\ \Sigma_{21} & \Sigma_2 \end{array}\right)\right).$$

If $\Sigma_2$ is invertible, the conditional density of the random vector $X_1$ given $X_2 = x_2$, is gaussian with mean $\mu_{1|2}$ and covariance $\Sigma_{1|2}$ such that

$$\mu_{1|2} = \mu_1 + \Sigma_{12}\Sigma_2^{-1}(x_2 - \mu_2)$$

and

$$\Sigma_{1|2} = \Sigma_1 - \Sigma_{12}\Sigma_2^{-1}\Sigma_{21}.$$