

MATH 545: INTRODUCTION TO TIME SERIES ANALYSIS

WINTER 2025

THE KALMAN FILTER RECURSIONS

Recall the simplest form of linear Gaussian state space model: for $t = 1, \dots, n$

$$\mathbf{Y}_t = \mathbf{F}_t \mathbf{X}_t + \mathbf{V}_t, \quad (1a)$$

$$\mathbf{X}_{t+1} = \mathbf{G}_t \mathbf{X}_t + \mathbf{W}_t. \quad (1b)$$

where

- $\mathbf{Y}_t$ is $(K \times 1)$,
- $\mathbf{X}_t$ is $(L \times 1)$,
- $\mathbf{V}_t \sim \text{Normal}_K(\mathbf{0}, \Sigma_t)$, mutually independent,
- $\mathbf{W}_t \sim \text{Normal}_L(\mathbf{0}, \Omega_t)$, mutually independent,
- $\mathbf{V}_s$ and $\mathbf{W}_t$ independent for all s, t .

For the Kalman filter recursions, let

- $\mathbf{a}_{t|t} = \mathbb{E}[\mathbf{X}_t | \mathbf{y}_{1:t}]$,
- $\mathbf{a}_{t+1|t} = \mathbb{E}[\mathbf{X}_{t+1} | \mathbf{y}_{1:t}]$,
- $\mathbf{P}_{t|t} = \text{MSE}[\mathbf{X}_t | \mathbf{y}_{1:t}] = \text{Var}[\mathbf{X}_t | \mathbf{y}_{1:t}]$,
- $\mathbf{P}_{t+1|t} = \text{MSE}[\mathbf{X}_{t+1} | \mathbf{y}_{1:t}] = \text{Var}[\mathbf{X}_{t+1} | \mathbf{y}_{1:t}]$.

If we let $\mathbf{X}_1 \sim \text{Normal}(\mathbf{a}_{1|0}, \mathbf{P}_{1|0})$, then under the model assumptions, the full joint distribution

$$p(\mathbf{x}_{1:n}, \mathbf{y}_{1:n})$$

is a $2n$ dimensional multivariate Gaussian distribution. Hence, by the earlier results, all the conditional distributions of the form $p(\mathbf{x}_t | \mathbf{y}_{1:s})$ and $p(\mathbf{y}_t | \mathbf{y}_{1:s})$ for $s \leq t$ are also Gaussian.

The objective is to compute the conditional expectations and variances in a recursive fashion.

- *Filtering*: compute $(\mathbf{a}_{t|t}, \mathbf{P}_{t|t})$ from $(\mathbf{a}_{t|t-1}, \mathbf{P}_{t|t-1})$
- *Prediction*: compute $(\mathbf{a}_{t+1|t}, \mathbf{P}_{t+1|t})$ from $(\mathbf{a}_{t|t}, \mathbf{P}_{t|t})$

Filtering: In the basic form of the model, filtering distribution is computed as follows:

$$p(\mathbf{x}_t | \mathbf{y}_{1:t}) \propto p(\mathbf{y}_t | \mathbf{x}_t) p(\mathbf{x}_t | \mathbf{y}_{1:(t-1)}) \quad (2)$$

Now, under observation and state equations,

$$p(\mathbf{x}_t | \mathbf{y}_{1:(t-1)}) \equiv \text{Normal}(\mathbf{a}_{t|(t-1)}, \mathbf{P}_{t|(t-1)}) \quad p(\mathbf{y}_t | \mathbf{x}_t) \equiv \text{Normal}(\mathbf{F}_t \mathbf{x}_t, \Sigma_t).$$

The right-hand side of (2), when viewed as a function of $\mathbf{x}_t$ is, up to proportionality,

$$\exp \left\{ -\frac{1}{2} \left[(\mathbf{y}_t - \mathbf{F}_t \mathbf{x}_t)^\top \Sigma_t^{-1} (\mathbf{y}_t - \mathbf{F}_t \mathbf{x}_t) + (\mathbf{x}_t - \mathbf{a}_{t|(t-1)})^\top \mathbf{P}_{t|(t-1)}^{-1} (\mathbf{x}_t - \mathbf{a}_{t|(t-1)}) \right] \right\}$$

The term in square brackets is a quadratic form in $\mathbf{x}_t$; writing this as

$$(\mathbf{x}_t - \mathbf{a}_{t|t})^\top \mathbf{P}_{t|t}^{-1} (\mathbf{x}_t - \mathbf{a}_{t|t}) + \text{constant}$$

we can compare terms in the two expressions. The quadratic term $\mathbf{x}_t \mathbf{M}_t^{-1} \mathbf{x}_t$ yields

$$\mathbf{P}_{t|t} = (\mathbf{F}_t^\top \boldsymbol{\Sigma}_t^{-1} \mathbf{F}_t + \mathbf{P}_{t|t-1}^{-1})^{-1}$$

and the linear term yields that

$$\mathbf{P}_{t|t}^{-1} \mathbf{a}_{t|t} = \mathbf{F}_t^\top \boldsymbol{\Sigma}_t^{-1} \mathbf{y}_t + \mathbf{P}_{t|t-1}^{-1} \mathbf{a}_{t|t-1}$$

so that

$$\mathbf{a}_{t|t} = (\mathbf{F}_t^\top \boldsymbol{\Sigma}_t^{-1} \mathbf{F}_t + \mathbf{P}_{t|t-1}^{-1})^{-1} (\mathbf{F}_t^\top \boldsymbol{\Sigma}_t^{-1} \mathbf{y}_t + \mathbf{P}_{t|t-1}^{-1} \mathbf{a}_{t|t-1})$$

Now, recall the Woodbury matrix formula: where $\mathbf{A}$ and $\mathbf{C}$ are square and non-singular, then

$$(\mathbf{A} + \mathbf{UCV})^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1} \mathbf{U} (\mathbf{C}^{-1} + \mathbf{VA}^{-1} \mathbf{U})^{-1} \mathbf{VA}^{-1}.$$

In this formula we identify

- $\mathbf{A}$ as $\mathbf{P}_{t|t-1}^{-1}$
- $\mathbf{C}$ as $\boldsymbol{\Sigma}_t^{-1}$
- $\mathbf{U}$ as $\mathbf{F}_t^\top$
- $\mathbf{V}$ as $\mathbf{F}_t$

and write

$$\mathbf{P}_{t|t} = \mathbf{P}_{t|t-1} - \mathbf{P}_{t|t-1} \mathbf{F}_t^\top (\boldsymbol{\Sigma}_t + \mathbf{F}_t \mathbf{P}_{t|t-1} \mathbf{F}_t^\top)^{-1} \mathbf{F}_t \mathbf{P}_{t|t-1}$$

If we define

$$\mathbf{M}_t = \boldsymbol{\Sigma}_t + \mathbf{F}_t \mathbf{P}_{t|t-1} \mathbf{F}_t^\top$$

then

$$\mathbf{P}_{t|t} = \mathbf{P}_{t|t-1} - \mathbf{P}_{t|t-1} \mathbf{F}_t^\top \mathbf{M}_t^{-1} \mathbf{F}_t \mathbf{P}_{t|t-1} \quad (3)$$

and

$$\begin{aligned} \mathbf{a}_{t|t} &= (\mathbf{P}_{t|t-1} - \mathbf{P}_{t|t-1} \mathbf{F}_t^\top \mathbf{M}_t^{-1} \mathbf{F}_t \mathbf{P}_{t|t-1}) (\mathbf{F}_t^\top \boldsymbol{\Sigma}_t^{-1} \mathbf{y}_t + \mathbf{P}_{t|t-1}^{-1} \mathbf{a}_{t|t-1}) \\ &= (\mathbf{I} - \mathbf{P}_{t|t-1} \mathbf{F}_t^\top \mathbf{M}_t^{-1} \mathbf{F}_t) (\mathbf{P}_{t|t-1} \mathbf{F}_t^\top \boldsymbol{\Sigma}_t^{-1} \mathbf{y}_t + \mathbf{a}_{t|t-1}) \\ &= \mathbf{P}_{t|t-1} \mathbf{F}_t^\top \boldsymbol{\Sigma}_t^{-1} \mathbf{y}_t + \mathbf{a}_{t|t-1} - \mathbf{P}_{t|t-1} \mathbf{F}_t^\top \mathbf{M}_t^{-1} \mathbf{F}_t \mathbf{P}_{t|t-1} \mathbf{F}_t^\top \boldsymbol{\Sigma}_t^{-1} \mathbf{y}_t - \mathbf{P}_{t|t-1} \mathbf{F}_t^\top \mathbf{M}_t^{-1} \mathbf{F}_t \mathbf{a}_{t|t-1} \\ &= \mathbf{a}_{t|t-1} + \mathbf{P}_{t|t-1} \mathbf{F}_t^\top (\boldsymbol{\Sigma}_t^{-1} \mathbf{y}_t - \mathbf{M}_t^{-1} \mathbf{F}_t \mathbf{P}_{t|t-1} \mathbf{F}_t^\top \boldsymbol{\Sigma}_t^{-1} \mathbf{y}_t - \mathbf{M}_t^{-1} \mathbf{F}_t \mathbf{a}_{t|t-1}) \\ &= \mathbf{a}_{t|t-1} + \mathbf{P}_{t|t-1} \mathbf{F}_t^\top (\boldsymbol{\Sigma}_t^{-1} \mathbf{y}_t - \mathbf{M}_t^{-1} (\mathbf{M}_t - \boldsymbol{\Sigma}_t) \boldsymbol{\Sigma}_t^{-1} \mathbf{y}_t - \mathbf{M}_t^{-1} \mathbf{F}_t \mathbf{a}_{t|t-1}) \\ &= \mathbf{a}_{t|t-1} + \mathbf{P}_{t|t-1} \mathbf{F}_t^\top (\boldsymbol{\Sigma}_t^{-1} \mathbf{y}_t - (\boldsymbol{\Sigma}_t^{-1} - \mathbf{M}_t^{-1}) \mathbf{y}_t - \mathbf{M}_t^{-1} \mathbf{F}_t \mathbf{a}_{t|t-1}) \\ &= \mathbf{a}_{t|t-1} + \mathbf{P}_{t|t-1} \mathbf{F}_t^\top \mathbf{M}_t^{-1} (\mathbf{y}_t - \mathbf{F}_t \mathbf{a}_{t|t-1}) \\ &= \mathbf{a}_{t|t-1} + \mathbf{P}_{t|t-1} \mathbf{F}_t^\top \mathbf{M}_t^{-1} \mathbf{r}_t \end{aligned} \quad (4)$$

which yields in (4) the filtering mean and in (3) the filtering variance. Thus, in this Gaussian case

$$\mathbf{X}_t | \mathbf{y}_{1:t} \sim \text{Normal}_L(\mathbf{a}_{t|t}, \mathbf{P}_{t|t})$$

Prediction: The prediction distribution for $\mathbf{X}_{t+1}$ given $\mathbf{Y}_{1:t}$ is computed as follows:

$$p(\mathbf{x}_{t+1}|\mathbf{y}_{1:t}) = \int p(\mathbf{x}_{t+1}|\mathbf{x}_t)p(\mathbf{x}_t|\mathbf{y}_{1:t}) d\mathbf{x}_t$$

We may compute this integral analytically. However, more simply, we may use the moment formulae, and iterated expectation and variance by exploiting the simple form of the state equation. For example

$$\mathbf{a}_{t+1|t} = \mathbb{E}[\mathbf{X}_{t+1}|\mathbf{y}_{1:t}] = \mathbb{E}_{\mathbf{X}_t|\mathbf{Y}_{1:t}}[\mathbb{E}[\mathbf{X}_{t+1}|\mathbf{X}_t, \mathbf{y}_{1:t}]|\mathbf{y}_{1:t}] = \mathbb{E}_{\mathbf{X}_t|\mathbf{Y}_{1:t}}[\mathbf{G}_t \mathbf{X}_t|\mathbf{y}_{1:t}] = \mathbf{G}_t \mathbf{a}_{t|t} \quad (5)$$

and also

$$\begin{aligned} \mathbf{P}_{t+1|t} = \text{Var}[\mathbf{X}_{t+1}|\mathbf{y}_{1:t}] &= \text{Var}_{\mathbf{X}_t|\mathbf{Y}_{1:t}}[\mathbb{E}[\mathbf{X}_{t+1}|\mathbf{X}_t, \mathbf{y}_{1:t}]|\mathbf{y}_{1:t}] + \mathbb{E}_{\mathbf{X}_t|\mathbf{Y}_{1:t}}[\text{Var}[\mathbf{X}_{t+1}|\mathbf{X}_t, \mathbf{y}_{1:t}]|\mathbf{y}_{1:t}] \\ &= \text{Var}_{\mathbf{X}_t|\mathbf{Y}_{1:t}}[\mathbf{G}_t \mathbf{X}_t|\mathbf{y}_{1:t}] + \mathbb{E}_{\mathbf{X}_t|\mathbf{Y}_{1:t}}[\mathbf{\Omega}_t|\mathbf{y}_{1:t}] \\ &= \mathbf{G}_t \mathbf{P}_{t|t} \mathbf{G}_t^\top + \mathbf{\Omega}_t \end{aligned} \quad (6)$$

Hence

$$\mathbf{X}_{t+1}|\mathbf{y}_{1:t} \sim \text{Normal}_L(\mathbf{G}_t \mathbf{a}_{t|t}, \mathbf{G}_t \mathbf{P}_{t|t} \mathbf{G}_t^\top + \mathbf{\Omega}_t)$$

Note that we may substitute the previous recursions (3) and (4) for $\mathbf{a}_{t|t}$ and $\mathbf{P}_{t|t}$ to get the direct recursions for $\mathbf{X}_t$, that is

$$\mathbf{a}_{t+1|t} = \mathbf{G}_t \mathbf{a}_{t|t} = \mathbf{G}_t(\mathbf{a}_{t|t-1} + \mathbf{P}_{t|t-1} \mathbf{F}_t^\top \mathbf{M}_t^{-1} \mathbf{r}_t)$$

and

$$\mathbf{P}_{t+1|t} = \mathbf{G}_t \mathbf{P}_{t|t} \mathbf{G}_t^\top + \mathbf{\Omega}_t = \mathbf{G}_t(\mathbf{P}_{t|t-1} - \mathbf{P}_{t|t-1} \mathbf{F}_t^\top \mathbf{M}_t^{-1} \mathbf{F}_t \mathbf{P}_{t|t-1}) \mathbf{G}_t^\top + \mathbf{\Omega}_t$$

Finally, the prediction distribution for $\mathbf{Y}_{t+1}$ given $\mathbf{Y}_{1:t}$ is computed as follows:

$$p(\mathbf{y}_{t+1}|\mathbf{y}_{1:t}) = \int p(\mathbf{y}_{t+1}|\mathbf{x}_{t+1})p(\mathbf{x}_{t+1}|\mathbf{y}_{1:t}) d\mathbf{x}_t$$

By a similar argument to above, we have by iterated expectation

$$\mathbb{E}[\mathbf{Y}_{t+1}|\mathbf{y}_{1:t}] = \mathbb{E}_{\mathbf{X}_{t+1}|\mathbf{Y}_{1:t}}[\mathbb{E}[\mathbf{Y}_{t+1}|\mathbf{X}_{t+1}, \mathbf{y}_{1:t}]|\mathbf{y}_{1:t}] = \mathbb{E}_{\mathbf{X}_{t+1}|\mathbf{Y}_{1:t}}[\mathbf{F}_{t+1} \mathbf{X}_{t+1}|\mathbf{y}_{1:t}] = \mathbf{F}_{t+1} \mathbf{a}_{t+1|t} \quad (7)$$

and also

$$\begin{aligned} \text{Var}[\mathbf{Y}_{t+1}|\mathbf{y}_{1:t}] &= \text{Var}_{\mathbf{X}_{t+1}|\mathbf{Y}_{1:t}}[\mathbb{E}[\mathbf{Y}_{t+1}|\mathbf{X}_{t+1}, \mathbf{y}_{1:t}]|\mathbf{y}_{1:t}] + \mathbb{E}_{\mathbf{X}_{t+1}|\mathbf{Y}_{1:t}}[\text{Var}[\mathbf{Y}_{t+1}|\mathbf{X}_{t+1}, \mathbf{y}_{1:t}]|\mathbf{y}_{1:t}] \\ &= \text{Var}_{\mathbf{X}_{t+1}|\mathbf{Y}_{1:t}}[\mathbf{F}_{t+1} \mathbf{X}_{t+1}|\mathbf{y}_{1:t}] + \mathbb{E}_{\mathbf{X}_{t+1}|\mathbf{Y}_{1:t}}[\mathbf{\Sigma}_t|\mathbf{y}_{1:t}] \\ &= \mathbf{F}_{t+1} \mathbf{P}_{t+1|t} \mathbf{F}_{t+1}^\top + \mathbf{\Sigma}_t \end{aligned} \quad (8)$$

Hence

$$\mathbf{Y}_{t+1}|\mathbf{y}_{1:t} \sim \text{Normal}_L(\mathbf{F}_{t+1} \mathbf{a}_{t+1|t}, \mathbf{F}_{t+1} \mathbf{P}_{t+1|t} \mathbf{F}_{t+1}^\top + \mathbf{\Sigma}_t)$$