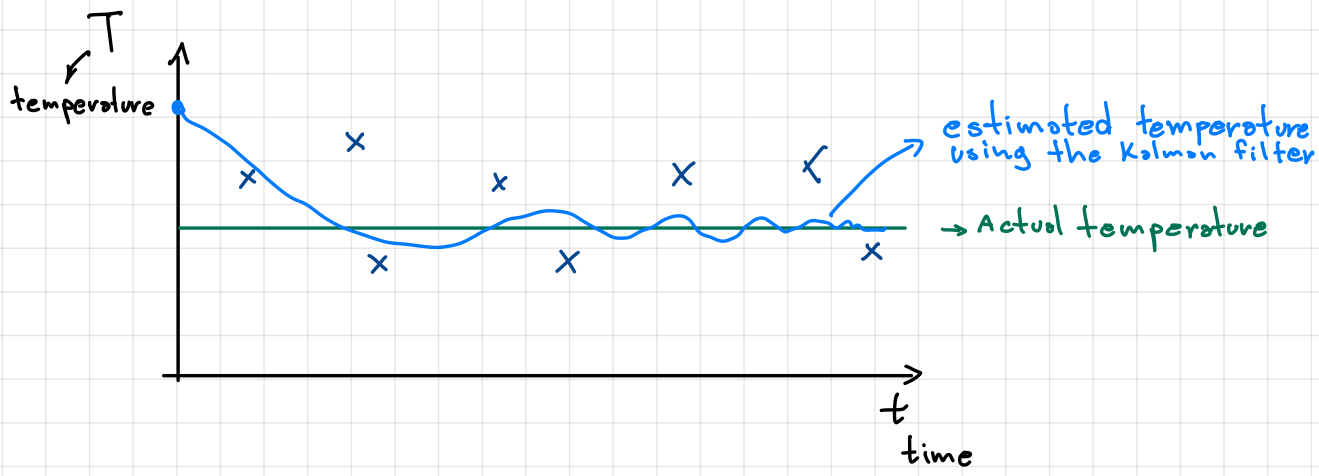
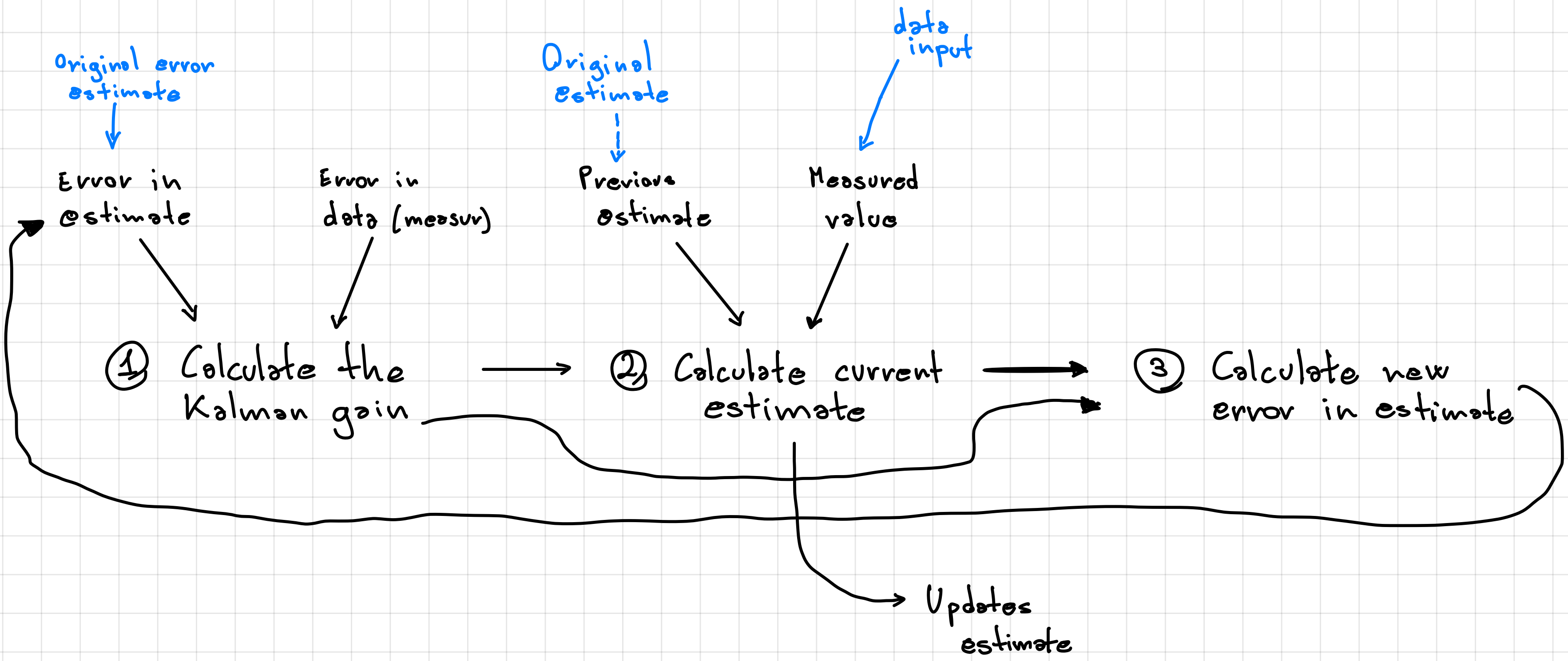
Kalman filter

- iterative mathematical process that uses a set of equations and consecutive data inputs to estimate the true value of an object when the measured values contain unpredicted/random error



Flowchart



• A closer look to the Kalman gain

Kalman gain: how much of the new measurement to use to update the new estimate

$$K = \frac{E_{est}}{E_{est} + E_{meas}}$$

$0 \leq K \leq 1$

E_{est} : error in estimate

E_{meas} : error in measurement

: see flowchart

if we have the

current estimate $\Rightarrow$ EST_t

previous estimate $\Rightarrow$ EST_{t-1}

measurement $\Rightarrow$ MEA

$$EST_t = EST_{t-1} + K \cdot \left[MEA - EST_{t-1} \right]$$

measurements are accurate	K	estimates are unstable
	1	
	0,8	
	0,2	
	0	
measurements are inaccurate		estimates are stable

① Calculate the Kalman gain $\longrightarrow$ ② Calculate the current estimate $\longrightarrow$ ③ Calculate the new est. error

1) $$K = \frac{E_{est}}{E_{est} + E_{meas}}$$

2) $$EST_t = EST_{t-1} + K \cdot [MEA - EST_{t-1}]$$

3) $$E_{est_t} = \frac{(E_{meas}) \cdot (E_{est_{t-1}})}{(E_{meas}) \cdot (E_{est_{t-1}})} \Longrightarrow E_{est_t} = [1 - K] \cdot (E_{est_{t-1}})$$

smaller than $E_{est_{t-1}}$!

• Example

Temperature:

- true temperature = 72
- initial estimate = 68
- initial $E_{est} = 2$
- init. measur. = 75
- error measur. = 4

first iteration

$$K = \frac{2}{2 + 4} = 0,33$$

$$EST_t = 68 + 0,33 \cdot [75 - 68] = 70,33$$

$$E_{est_t} = [1 - 0,33] \cdot 2 = 1,33$$

• The multi-dimension model

initial state: $\boxed{\begin{matrix} X_0 \\ P_0 \end{matrix}}$ - - - - - > Previous state: X_{t-1}, P_{t-1} - - - > New state (predicted)

$$X_{t_p} = A X_{t-1} + B \cdot u_t + w_t$$

$$P_{t_p} = A \cdot P_{t-1} \cdot A^T + Q_t$$

• Statistical introduction

Take $\alpha_{t+1|t} = E(\alpha_{t+1} | \mathcal{Y}_t)$ (information set) — cond. exp. of the state given the past

$P_{t+1|t} = V(\alpha_{t+1} | \mathcal{Y}_t)$ — cond. variance of the unobs. state given the past

and the initial condition as

$$\left.\begin{aligned} \alpha_{1|0} &= \alpha_1 \\ P_{1|0} &= P_1 \end{aligned}\right\} \text{given}$$

Let's now consider the prediction error

$$\begin{aligned}
v_t &= y_t - E[y_t | \mathcal{Y}_{t-1}] \quad \text{best MMSE predictor of } y_t \text{ given the past} \\
&= y_t - E[Z_t \cdot \alpha_t + G_t \cdot \varepsilon_t | \mathcal{Y}_{t-1}] \quad \text{from the obs. equation } y_t = Z_t \cdot \alpha_t + G_t \cdot \varepsilon_t \\
&= y_t - Z_t \cdot E(\alpha_t | \mathcal{Y}_{t-1}) \quad \text{by linearity of cond. exp } E(G_t \cdot \varepsilon_t | \mathcal{Y}_{t-1}) = G_t \cdot E(\varepsilon_t | \mathcal{Y}_{t-1}) = 0 \\
&= y_t - Z_t \cdot \alpha_{t|t-1} \\
&= (Z_t \cdot \alpha_t + G_t \cdot \varepsilon_t - Z_t \cdot \alpha_{t|t-1}) \\
&= Z_t (\alpha_t - \alpha_{t|t-1}) + G_t \cdot \varepsilon_t
\end{aligned}$$

we have

$$E(v_t | \mathcal{Y}_{t-1}) = 0$$

$$V(v_t | \mathcal{Y}_{t-1}) = Z_t \cdot P_{t|t-1} \cdot Z_t' + G_t \cdot G_t'$$

the idea:

• Now, consider the updating:

in $t-1$, our best (MMSE) prediction of y_t given the past information set $\mathcal{Y}_{t-1}$ is $Z_t \cdot \alpha_{t|t-1}$. $= E[y_t | \mathcal{Y}_{t-1}]$

Then the actual observation y_t arrives we compute the prediction error $v_t = y_t - Z_t \cdot \alpha_{t|t-1}$ and its variance $F_t = Z_t \cdot P_{t|t-1} \cdot Z_t' + G_t \cdot G_t'$.

The new best estimate of α_t is then updated and now based on the old estimate $\alpha_{t|t-1}$ and the new information included by v_t:

$$\alpha_{t+1|t} = T_t \cdot \alpha_{t|t-1} + K_t \cdot v_t$$

and for the variance:

$$P_{t+1|t} = T_t \cdot P_t \cdot T_t' + H_t \cdot H_t' - K_t \cdot F_t \cdot K_t'$$

where

$$K_t = T_t \cdot P_{t|t-1} \cdot Z_t' \cdot F_t^{-1}$$

• Equations of the Kalman filter

- $v_t = y_t - Z_t \cdot \alpha_{t|t-1}$ → measures the mismatch between the observed data and the model prediction

- $F_t = \underbrace{Z_t}_{\text{observation matrix}} \cdot \underbrace{P_{t|t-1}}_{\text{predict. error cov. matrix}} \cdot Z_t' + \underbrace{G_t \cdot G_t'}_{\text{noise covariance matrix}}$ → compute the innovation covariance matrix, quantifies the uncertainty or variance in the predicted measurements

- $K_t = \underbrace{T_t}_{\text{transition matrix}} \cdot P_{t|t-1} \cdot Z_t' \cdot F_t^{-1}$ → Kalman gain: the amount of weight given to the observed measurement in updating the state estimate

- $P_{t+1|t} = T_t \cdot P_t \cdot T_t' + H_t \cdot \underbrace{Q_t}_{\text{noise covariance matrix}} \cdot H_t' - K_t \cdot F_t \cdot K_t'$ → updates the error cov. matrix P based on the measurement

- $\alpha_{t+1|t} = T_t \cdot \alpha_{t|t-1} + K_t \cdot v_t$ → updates the state estimate α based on the measurement

In summary, the Kalman filter equations allow for the recursive estimation of the hidden state α by combining the predicted state with the observed measurement. The filter incorporates the system dynamics, measurement uncertainties, and the relationship between the predicted and observed measurements to provide an optimal estimate of the hidden state at each time step.

So with the recursion at time 1 we have:

$$V_1 = y_1 - Z \cdot \alpha_{1|0} = y_1 - Z \cdot \alpha_1$$

$$F_1 = Z \cdot P_1 \cdot Z^T + G \cdot G^T$$

$$K_1 = T \cdot P_1 \cdot Z^T \cdot F_1^{-1}$$

$$\alpha_{2|1} = T \cdot \alpha_{1|0} + K_1 V_1$$

$$P_{2|1} = T \cdot P_1 \cdot T^T + H \cdot Q \cdot H^T - K_1 \cdot F_1 \cdot K_1$$

• The Kalman filter proof

The derivation of the Kalman filter can be regarded as an application of a regression lemma for the multi-variate normal distribution

• Lemma

proven proposition used as a stepping stone to a larger result

Let x, y, z be jointly Gaussian with $E(z) = 0$ and $\Sigma_{yz} = 0$.

Then

$$E(x \mid y, z) = E(x \mid y) + \Sigma_{xz} \cdot \Sigma_{zz}^{-1} \cdot z$$

$$Var(x \mid y, z) = Var(x \mid y) - \Sigma_{x \cdot z} \cdot \Sigma_{zz}^{-1} \cdot \Sigma_{xz}'$$

proof:

Let $\mathcal{Y}_t = \sigma\{y_t, y_{t-1}, \dots\} = \sigma\{v_t, y_{t-1}, \dots\}$

and $E(v_t \cdot y_{t-j}) = 0$ for $j > 0$ as v_t is a martingale difference sequence and

$E\left(v_t \cdot g(\mathcal{Y}_{t-j})\right) = 0$ for $j > 0$

Recall the lemma $E(x \mid y, z) = E(x \mid y) + \Sigma_{xz} \cdot \Sigma_{zz}^{-1} \cdot z$ and set

$$x = \alpha_{t+1}$$

$$y = \mathcal{Y}_{t-1} \quad (y_{t-1})$$

$$z = v_t = Z_t(\alpha_t - \alpha_{t|t-1}) + G_t \cdot \varepsilon_t$$

and observe that from the transition equation $\rightarrow \alpha_{t+1} = T \cdot \alpha_t + H \cdot \eta_t$

$$E(\alpha_{t+1} \mid \mathcal{Y}_{t-1}) = T_t \cdot \alpha_{t|t-1}$$

So combining the last two equation

$$E(\alpha_{t+1} \cdot v_t' \mid \mathcal{Y}_{t-1}) = T_t \cdot E(\alpha_t \cdot v_t' \mid \mathcal{Y}_{t-1}) + H_t \cdot E(\eta_t \cdot v_t' \mid \mathcal{Y}_{t-1}) =$$

$$= T_t \cdot P_{t|t-1} \cdot Z_t'$$

From the lemma

$$E(\alpha_{t+1} \mid \mathcal{Y}_{t-1}, v_t) = T_t \cdot \alpha_{t|t-1} + T_t \cdot P_{t|t-1} \cdot Z_t' \cdot F_t^{-1} \cdot v_t$$

$$= T_t \cdot \alpha_{t|t-1} + K_t \cdot v_t$$

where $\underbrace{K_t}_{\text{Kalman Gain}} = T_t \cdot P_{t|t-1} \cdot Z_t' \cdot F_t^{-1}$ and

$$Var\left(\alpha_{t+1} \mid \mathcal{Y}_{t-1}, v_t\right) = P_{t+1|t}$$

with, from the transition equation and from the lemma,

$$\begin{aligned} P_{t+1|t} &= T_t \cdot P_{t|t-1} \cdot T_t' + H_t \cdot Q_t \cdot H_t' - T_t \cdot P_{t|t-1} \cdot Z_t' \cdot F_t^{-1} \cdot \left(T_t \cdot P_{t|t-1} \cdot Z_t'\right)' \\ &= T_t \cdot P_{t|t-1} \cdot T_t' + H_t \cdot Q_t \cdot H_t' - \left(T_t \cdot P_{t|t-1} \cdot Z_t' \cdot F_t^{-1}\right) \cdot F_t \left(F_t^{-1} \cdot Z_t \cdot P_{t|t-1} \cdot T_t'\right) \\ &= T_t \cdot P_{t|t-1} \cdot T_t' + H_t \cdot Q_t \cdot H_t' - K_t \cdot F_t^{-1} \cdot K_t' \end{aligned}$$

The Kalman filter updates the prediction of y_t given $\mathcal{Y}_{t-1}$ by calculating the mean and variance unconditional and conditional on y_t or v_t

We move by updates/estimate based on $\mathcal{Y}_{t-1}$ to updated estimates based on y_t or equiv. Y_t

The innovation form of Kalman filter

Let's recall what we have seen so far

1) UC model

$$y_t = \mu_t + \varepsilon_t \quad , \quad \varepsilon_t \sim N(0, \sigma_\varepsilon^2)$$

$$\mu_{t+1} = \phi \mu_t + \eta_t \quad , \quad \eta_t \sim N(0, \sigma_\eta^2)$$

2) State space representation

$$y_t = Z_t \cdot \alpha_t + G_t \cdot \varepsilon_t \quad , \quad \varepsilon_t \sim N(0, I)$$

$$\alpha_{t+1} = T_t \cdot \alpha_t + H_t \cdot \eta_t \quad , \quad \eta_t \sim N(0, Q_t)$$

3) Kalman filter

$$v_t = y_t - Z_t \cdot \alpha_{t|t-1}$$

$$\vdots$$

$$\alpha_{t+1|t} = T_t \cdot \alpha_{t|t-1} + K_t \cdot v_t$$

4) Innovation form of the Kalman filter

$$y_t = \mu_{t|t-1} + v_t$$

$$\mu_{t+1|t} = \phi \mu_{t|t-1} + K_t \cdot v_t$$

- in 1) μ_t is a function of μ_{t-1} and η_t is a r.v. with its own distr.
- in 4) $\mu_{t|t-1}$ is a function of $\mu_{t-1|t-2}$ and of the past observations $y_t, y_{t-1}, \dots$ incorporated through
  - the recursion
  - the prediction error
- in 1) we have a parameter driven form
- in 4) we have an observation driven representation

So, once we run the Kalman filter we can write the innovation form of the Kalman filter

• Kf in the LLM

Let's take the state space form of the LLM

• $Z_t = 1$, $\alpha_t = \mu_t$, $G_t = \sigma_\varepsilon$, $T_t = 1$, $H_t = \sigma_\eta$, $Q_t = 1$

and apply the Kalman filter recursion

$$v_t = y_t - Z_t \cdot \alpha_{t|t-1}$$

$$\alpha_{t+1|t} = T_t \cdot \alpha_{t|t-1} + K_t \cdot v_t$$

$$F_t = Z_t \cdot P_{t|t-1} \cdot Z_t' + G_t \cdot G_t'$$

$$K_t = T_t \cdot P_{t|t-1} \cdot Z_t' \cdot F_t^{-1}$$

$$P_{t+1|t} = T_t \cdot P_{t|t-1} \cdot T_t' + H_t \cdot Q_t \cdot H_t' - K_t \cdot F_t \cdot K_t'$$

So this leads to have

$$v_t = y_t - \mu_{t|t-1}$$

$$F_t = P_{t|t-1} + \sigma_\varepsilon^2$$

$$K_t = P_{t|t-1} / (P_{t|t-1} + \sigma_\varepsilon^2)$$

$$P_{t+1|t} = P_{t|t-1} + \sigma_\eta^2 - \left(\frac{P_{t|t-1}}{P_{t|t-1} + \sigma_\varepsilon^2}\right)^2 \cdot (P_{t|t-1} + \sigma_\varepsilon^2)$$

$$\mu_{t+1|t} = \mu_{t|t-1} + K_t \cdot v_t$$

with $\mu_{1|0}$ and $P_{1|0}$ fixed (to start the recursion)

$\mu_{t+1|t}$ emphasise the fact that $\mu_{t+1|t}$ is $\mathcal{Y}_t$-measurable (known given t)

• We can write it as a function of the Kalman gain $K_t = \frac{P_t}{F_t}$

$$v_t = y_t - \mu_{t|t-1}$$

$$F_t = P_{t|t-1} + \sigma_\varepsilon^2$$

$$P_{t+1|t} = P_{t|t-1} (1 - K_t) + \sigma_\eta^2$$

$$\mu_{t+1|t} = \mu_{t|t-1} + K_t \cdot v_t$$

with $\mu_{1|0}$ and $P_{1|0}$ fixed

The IFKF of an AR(1) + noise

$$y_t = \mu_{t|t-1} + v_t$$

$$\mu_{t+1|t} = \phi\, \mu_{t|t-1} + k_t \cdot v_t$$

which is equivalent to say that

$$y_t \mid \mathcal{Y}_{t-1} \sim \underbrace{N\left(\mu_{t|t-1}\,,\ F_t\right)}_{\text{conditional density}}$$