Kalman Filter

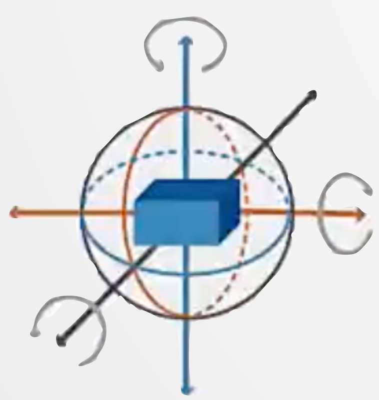
Ruting Wang

IRTG 1792
High Dimensional Nonstationary Time Series
Humboldt-Universität zu Berlin
IRTG1792.HU-Berlin.de

Not all variables (features) are observable!

- Variable of interest measured indirectly
  - Position

- Measurements might be subjected to noise
  - Acceleration
  - Angular velocity
  - Relative position



Inertial Measurement Unit (IMU)

Odometer

Applications

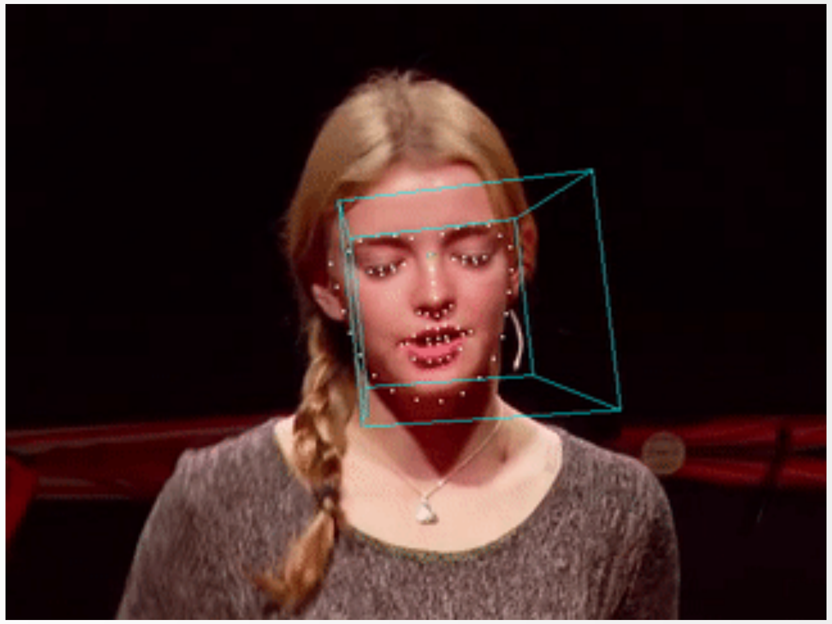
Navigation

Satellite receiver information

⋎

Location

Image Identification



Past photo information

⋎

Facial centre in the next frame

APOLLO and Kalman

“[…] not because *[it is]* easy, but because *[it is]* hard; because that goal will serve to organise and measure the best of our energies and skills, because that change is one that we are willing to accept, one we are unwilling to postpone, and one we intend to win […].”

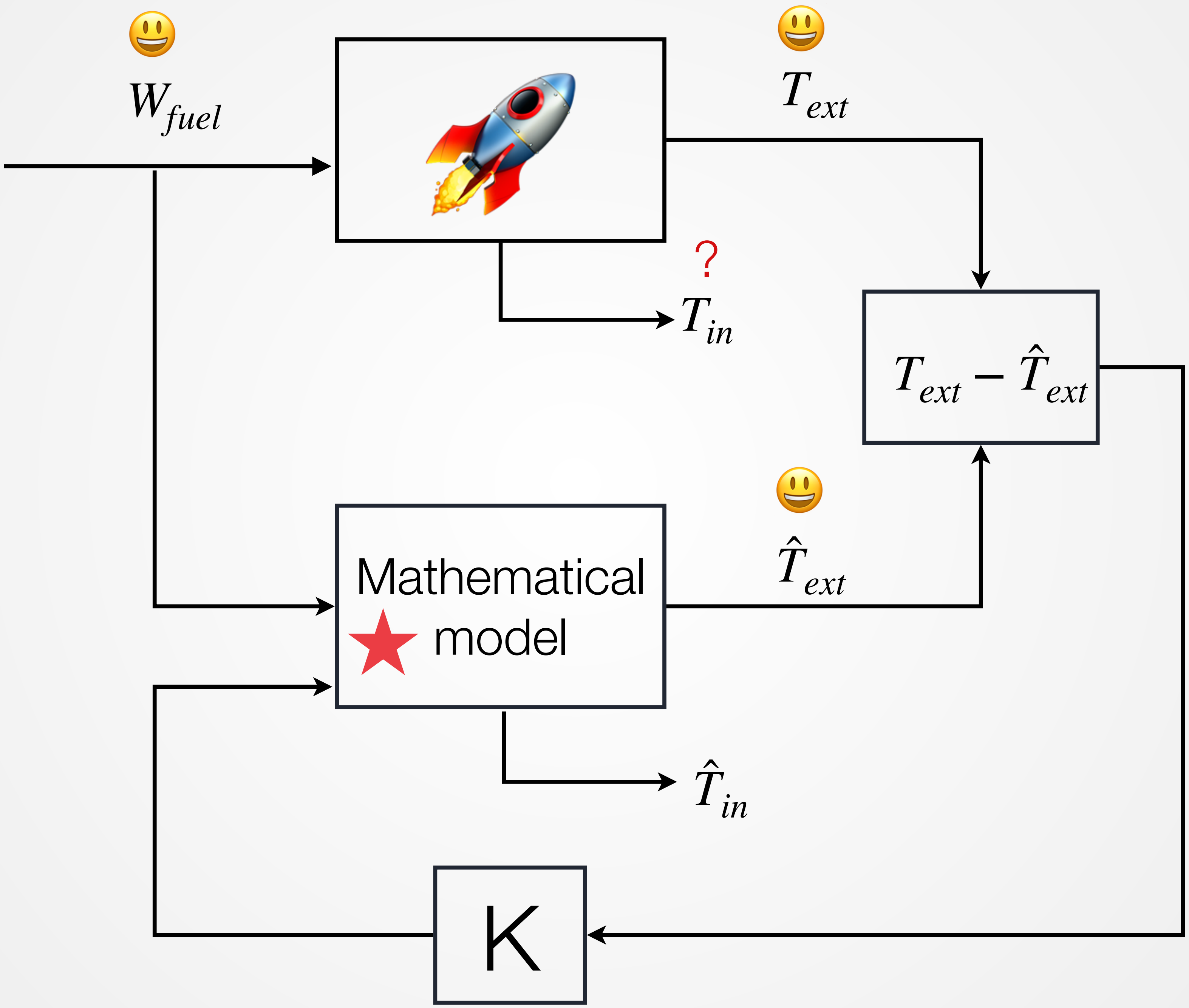
How about the Temperature inside APOLLO?

- Approximate what we not observe
  - Temperature inside the spaceship: T_{in}
  - Temperature in extreme situation
- Observable features 😃
  - Amount of fuel: W_{fuel}
  - Temperature outside the spaceship: T_{ext}
  - Mathematical model:

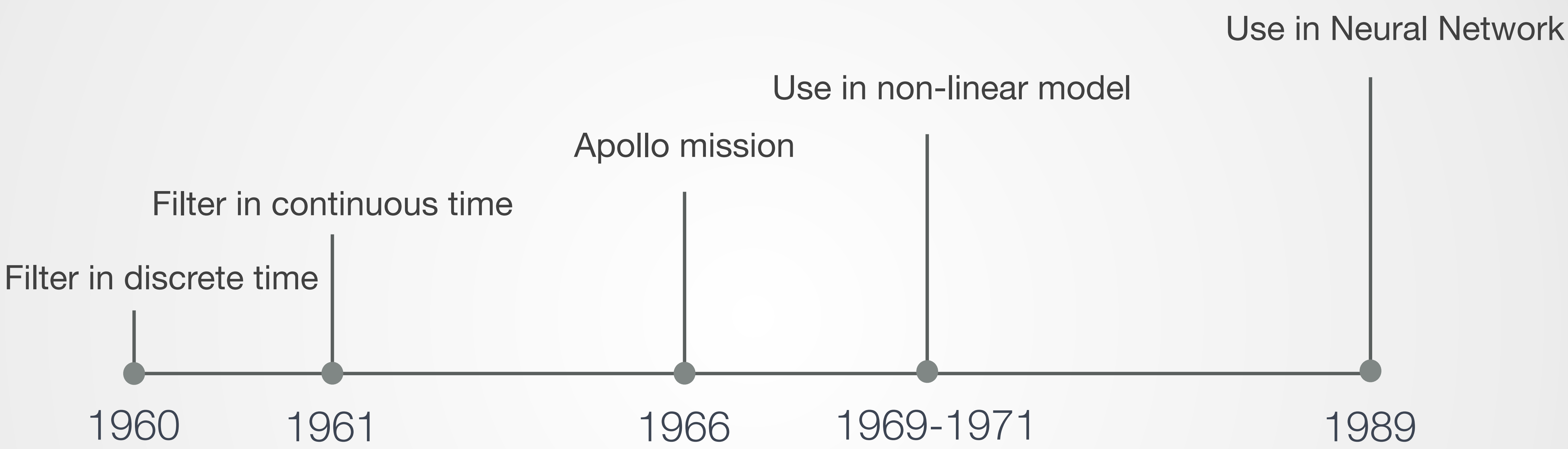
$$\hat{T}_{in} = B\ W_{fuel}$$

$$\hat{T}_{ext} = C\ \hat{T}_{in}$$

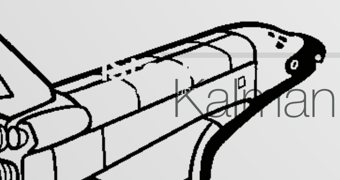
How to Measure the Temperature inside the Spaceship?



Kalman Filter Evolution



Rudolf Emil Kálmán on BBI



Outline

1. Motivation ✓
2. Algorithm Introduction
3. Application in Finance
4. References

A Simple Kalman Filter Case

- How to estimate a coin's real diameter x_t
  - The ruler's measurement results:

$$y_1 = 51\ mm$$
$$y_2 = 48\ mm$$
$$y_3 = 49\ mm$$
$$\cdots\cdot$$

  - Taking the average of y_t to get the estimation $\hat{x}_t$

$$\begin{aligned}
\hat{x}_t &= \frac{1}{t}(y_1 + y_2 + \ldots y_t) \\
&= \frac{1}{t}(y_1 + y_2 + \ldots y_{t-1}) + \frac{1}{t}y_t \\
&= \frac{1}{t}\frac{t-1}{t-1}(y_1 + y_2 + \ldots y_{t-1}) + \frac{1}{t}y_t \\
&= \frac{t-1}{t}\hat{x}_{t-1} + \frac{1}{t}y_t = \boxed{\hat{x}_{t-1} + \frac{1}{t}(y_t - \hat{x}_{t-1})}
\end{aligned}$$

Recursive idea!

A Simple Kalman Filter Case

- Kalman filter: also based on „Recursive“

$$\hat{x}_t = \hat{x}_{t-1} + \boxed{K_t}(y_t - \hat{x}_{t-1})$$

Kalman gain

- Estimation error: $e_t^E = x - \hat{x}_t$
- Measurement error: $e_t^M = y_t - x$

$$K_t = \frac{e_{t-1}^E}{e_{t-1}^E + e_t^M}$$

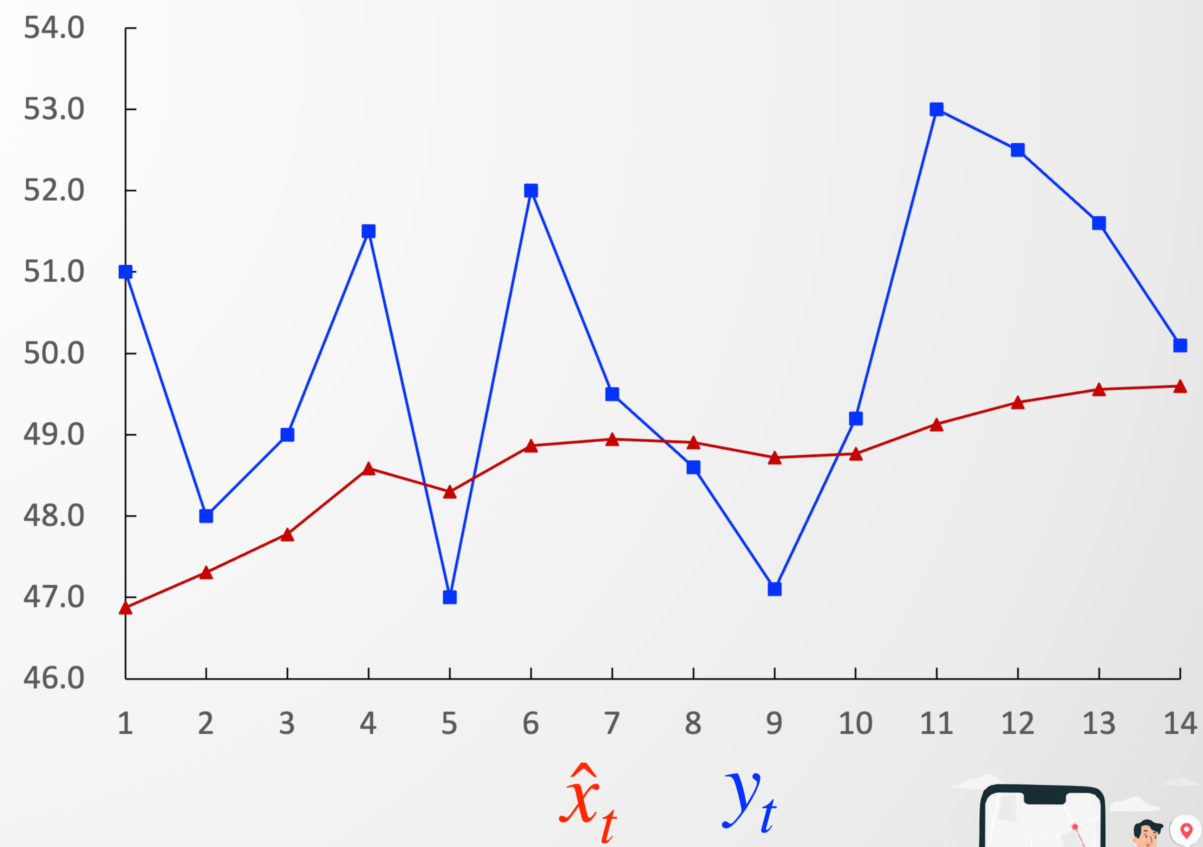
- $e_{t-1}^E \gg e_t^M \succ K_t = 1$ and $\hat{x}_t = y_t$
- $e_{t-1}^E \ll e_t^M \succ K_t = 0$ and $\hat{x}_t = \hat{x}_{t-1}$

A Simple Kalman Filter Case

- Step 1: calculate Kalman gain $K_t = \dfrac{e_{t-1}^E}{e_{t-1}^E + e_t^M}$
- Step 2: calculate $\hat{x}_t = \hat{x}_{t-1} + K_t(y_t - \hat{x}_{t-1})$
- Step 3: update $e_t^E = (1-K_t)e_{t-1}^E$

Initial input : $\hat{x}_0$ e_0^E Real value : $x = 50\ mm$

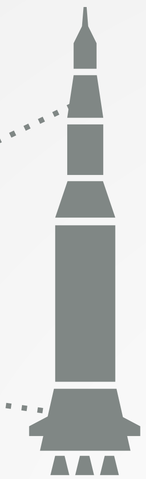
t	y	e^M	$\hat{x}_t$	K	e^E
0			40.000		5.000
1	51.0	3	46.875	0.625	1.875
2	48.0	3	47.308	0.385	1.154
3	49.0	3	47.778	0.278	0.833
4	51.5	3	48.587	0.217	0.652
5	47.0	3	48.304	0.179	0.536
6	52.0	3	48.864	0.152	0.455
7	49.5	3	48.947	0.132	0.395
8	48.6	3	48.907	0.116	0.349
9	47.1	3	48.719	0.104	0.313
10	49.2	3	48.764	0.094	0.283
11	53.0	3	49.129	0.086	0.259
12	52.5	3	49.397	0.079	0.238
13	51.6	3	49.559	0.074	0.221
14	50.1	3	49.596	0.068	0.205



A Complex APOLLO Case

- x is a Matrix

$$x = \begin{bmatrix} x^1 \\ x^2 \end{bmatrix}$$



- State function

$$x_t = Ax_{t-1} + Bu_{t-1} + w_{t-1} \quad \text{Process noise} \qquad w_{t-1} \sim N(0,Q)$$

$$y_t = Cx_t + v_t \quad \text{Measurement noise} \qquad v_t \sim N(0,R)$$

- Priori estimation

$$\hat{x}_t^- = A\hat{x}_{t-1} + Bu_{t-1}$$
$$\hat{x}_t^{mea} = C^{-1}y_t$$

- Posteriori estimation

$$\hat{x}_t = \hat{x}_t^- + K_t(y_k - C\hat{x}_t^-)$$

A Complex APOLLO Case

- Aim of Kalman filter: find a $K_t \succ min \left| x_t - \hat{x}_t \right|$

- Estimation error $e_t^E = x_t - \hat{x}_t$

- Properties of e_t^E

$$e^E \sim \mathrm{N}(0,P)$$

$$P = \mathrm{E}[e^E e^{E\top}] = \begin{bmatrix} \boxed{\sigma^2 e^{E1}} & \sigma e^{E1} \sigma e^{E2} \\ \sigma e^{E1} \sigma e^{E2} & \boxed{\sigma^2 e^{E2}} \end{bmatrix}$$

- $\boxed{\text{Find a } K_t \succ \min\{tr(P)\}}$

A Complex APOLLO Case

- Process of K's determination

$$e_t^E = x_t - \hat{x}_t = x_t - \hat{x}_t^- + K_t(y_t - C\hat{x}_t^-)$$

$$y_t = Cx_t + v_t$$

$$= (I - K_tC)\underbrace{(x_t - \hat{x}_t^-)}_{e_t^{E-}} - K_t v_t$$

$$P_t = \mathrm{E}[e^E e^{E\top}]$$

$$= (I - K_tC)\,\underbrace{\mathrm{E}\{e_t^{E-}(e_t^{E-})^\top\}}_{P_t^-}(I - K_tC)^\top + K_t\,\underbrace{\mathrm{E}(v_t v_t^\top)}_{R}K_t^\top$$

$$= P_t^- - K_tCP_t^- - P_t^-C^\top K_t^\top + K_tCP_t^-C^\top K_t^\top + K_tRK_t^\top$$

A Complex APOLLO Case

- Minimise the trace of P_t

$$tr(P_t) = tr(P_t^-) - 2tr(K_t C P_t^-) + tr(K_t C P_t^- C^\top K_t^\top) + tr(K_t R K_t^\top)$$

- Calculate Kalman Gain

$$\frac{\partial tr(P_t)}{\partial K_t} = 0 \quad \succ \quad \boxed{K_t = \frac{\hat{P}_t^- C^\top}{C\hat{P}_t^- C^\top + R}}$$

Kalman Gain!

A Complex APOLLO Case

- Calculate Priori P_t^-

$$x_t = Ax_{t-1} + Bu_{t-1} + w_{t-1}$$
$$\hat{x}_t^- = A\hat{x}_{t-1} + Bu_{t-1}$$

$\succ$

$$e_t^{E-} = x_t - \hat{x}_t^-$$
$$= Ax_{t-1} + Bu_{t-1} + w_{t-1} - A\hat{x}_{t-1} - Bu_{t-1}$$
$$= Ae_{t-1}^E + w_{t-1}$$

$$P_t^- = \mathrm{E}[e^{E-}e^{E-\top}]$$
$$= A\,\mathrm{E}[e_{t-1}^E e_{t-1}^{E\top}]A^\top + \mathrm{E}[w_{t_1} w_{t-1}^\top]$$
$$= AP_{t-1}A^\top + Q$$

- Update P_t^- with K_t $\succ$ Posteriori P_t

$$P_t = P_t^- - K_tCP_t^- - P_t^-C^\top K_t^\top + K_tCP_t^-C^\top K_t^\top + K_tRK_t^\top$$
$$= (I - K_tC)P_t^-$$

Estimation Process

- Prediction based on math model:

$$\hat{x}_t^- = A\hat{x}_{t-1} + Bu_t$$

$$\hat{P}_t^- = A\hat{P}_{t-1}A^\top + Q$$

$$\hat{y}_t^- = C\hat{x}_t^-$$

Without adjustment of y_t and Kalman gain K_t

Estimation Process

- Update process:

$$K_t = \frac{\hat{P}_t^- C^\top}{C\hat{P}_t^- C^\top + R}$$

$$\hat{x}_t = \hat{x}_t^- + K_t v_t = \hat{x}_t^- + K_t(y_t - C\hat{x}_t^-)$$

$$\hat{P}_t = (I - K_t C)\hat{P}_t^-$$

- Back to prediction process to estimate the state at *t+1*:

$$\hat{x}_{t+1}^- = A\hat{x}_t + Bu_t$$

$$\hat{P}_{t+1}^- = A\hat{P}_t A^\top + Q$$

$$\hat{y}_{t+1}^- = C\hat{x}_{t+1}^-$$

It is an application of Bayes
http://ais.informatik.uni-freiburg.de/teaching/ss10/robotics/slides/10-kalman-filter.pdf

Extreme Cases of Kalman Filter

- No measurement error

$$\lim_{R\to 0} K_t = \lim_{R\to 0} \frac{P_t^- C^\top}{CP_t^- C^\top + R} = \lim_{R\to 0} \frac{P_t^- C^\top}{CP_t^- C^\top + 0} = \frac{1}{C}$$

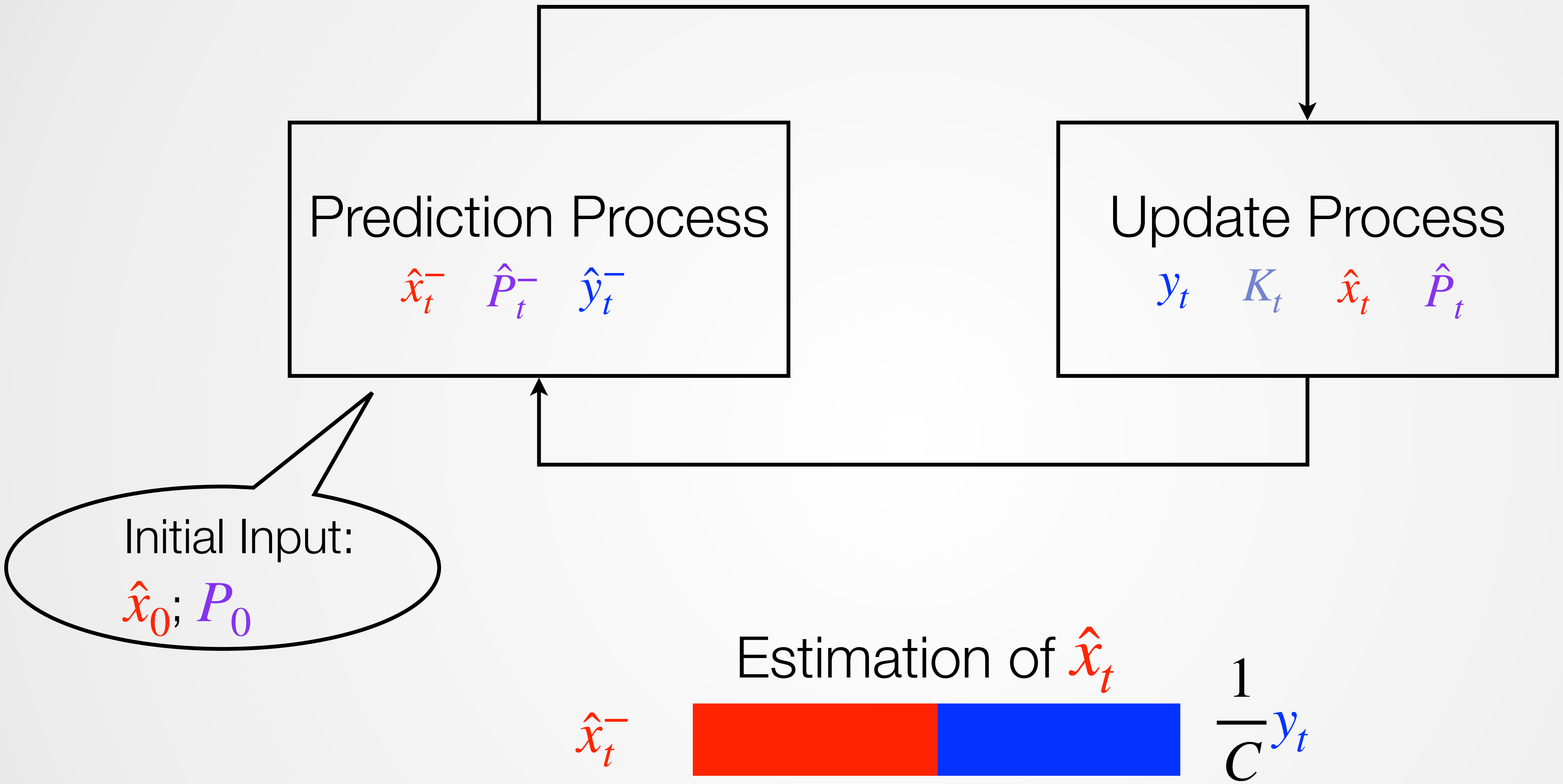
$$\hat{x}_t = \frac{1}{C} y_t$$

- No estimation error

$$\lim_{P_t^-\to 0} K_t = \lim_{P_t^-\to 0} \frac{P_t^- C^\top}{CP_t^- C^\top + R} = 0$$

$$\hat{x}_t = \hat{x}_t^-$$

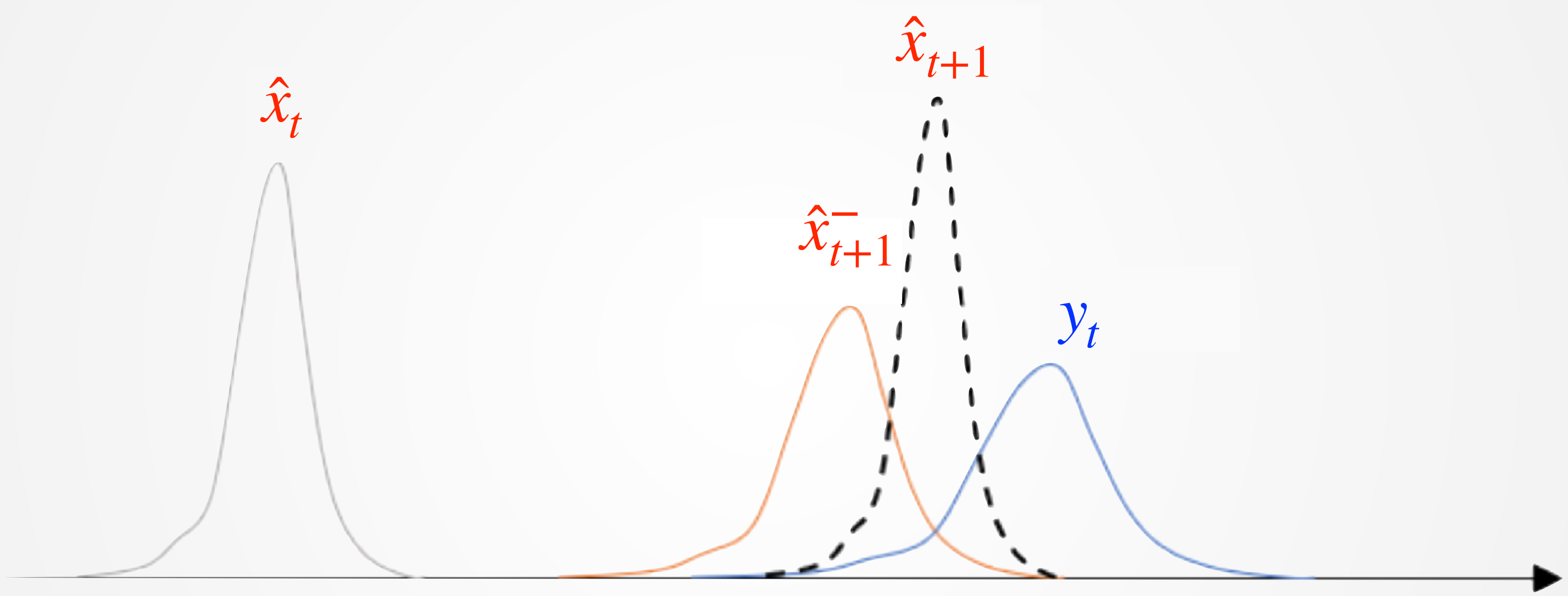
Dynamic Calculation Process



Framework

Estimation Process

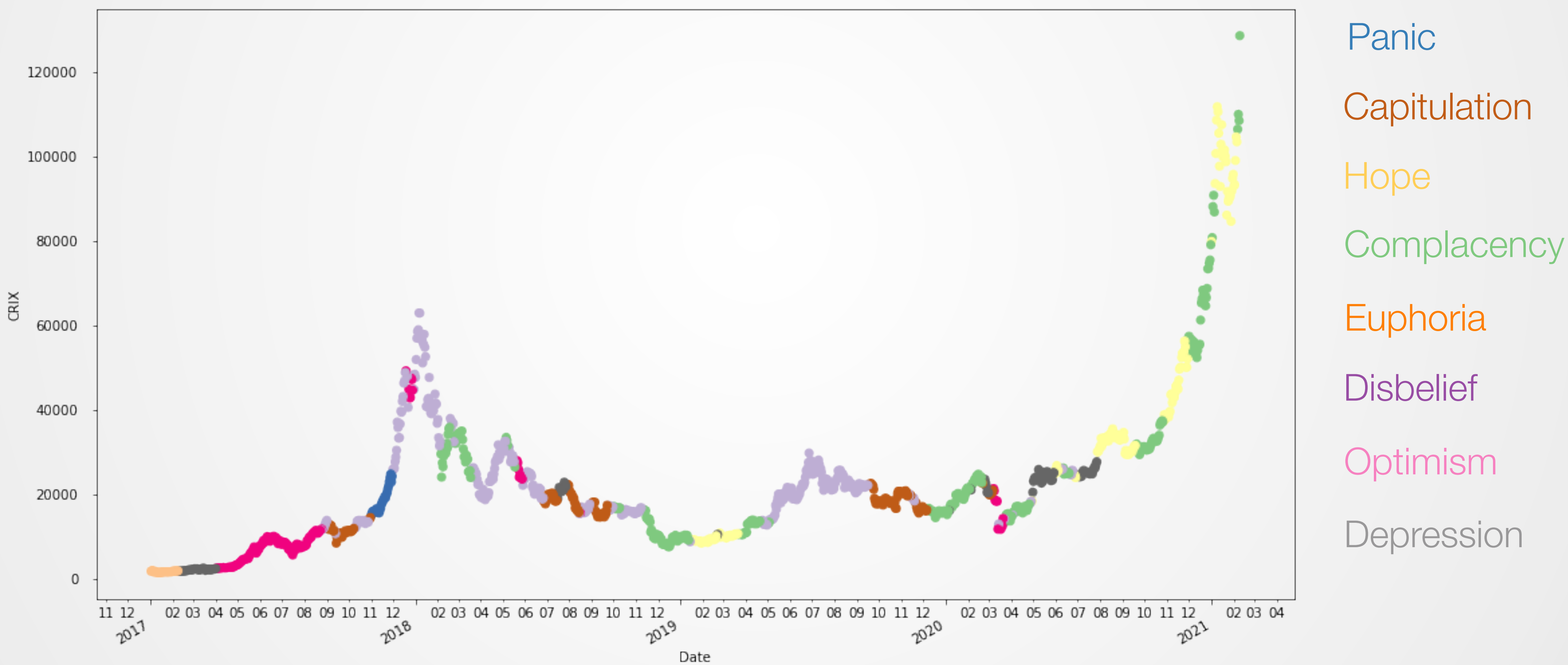
- Combine error distributions to get precise estimations



Dynamic Changes of Error Distribution

CRIX Return Prediction KF_modelres

- Kalman filter ≻ Return prediction
  - CRIX : A Laspeyre weighted index of market capitalisation.
  - Span: 2017-01-02 to 2021-02-09



CRIX cluster

Estimation Process

- Prediction based on math model:

$$\hat{r}^-_{x,t} = A\hat{r}_{x,t-1}$$

$$\hat{P}^-_t = A\hat{P}_{t-1}A^\top + Q$$

$$\hat{r}^-_{y,t} = C\hat{r}^-_{x,t}$$

$\hat{r}^-_{x,t}$: Predicted return of next day without adjustment

$\hat{r}^-_{y,t}$: return of today calculated by model

$Q : w_t \sim \mathrm{N}(0, 0.03)$

$A = 1 \;\; C = 1$

$\hat{r}_{x,0}$: The observable return on 2017-01-02

$\hat{P}_0 = 1$

Estimation Process

- Update Process:

$$K_t = \frac{\hat{P}_t^-}{\hat{P}_t^- + R}$$

$$\hat{r}_{x,t} = \hat{r}_{x,t}^- + K_t(r_{y,t} - \hat{r}_{x,t}^-)$$

$$\hat{P}_t = (I - K_t)\hat{P}_t^-$$

$\hat{r}_{x,t}$: Predicted return after adjustment

$\hat{r}_{y,t}$: return on day t that we observe

R : $v_t \sim \mathrm{N}(0, 0.03)$,

- Back to prediction process to estimate the state at *t+1*:

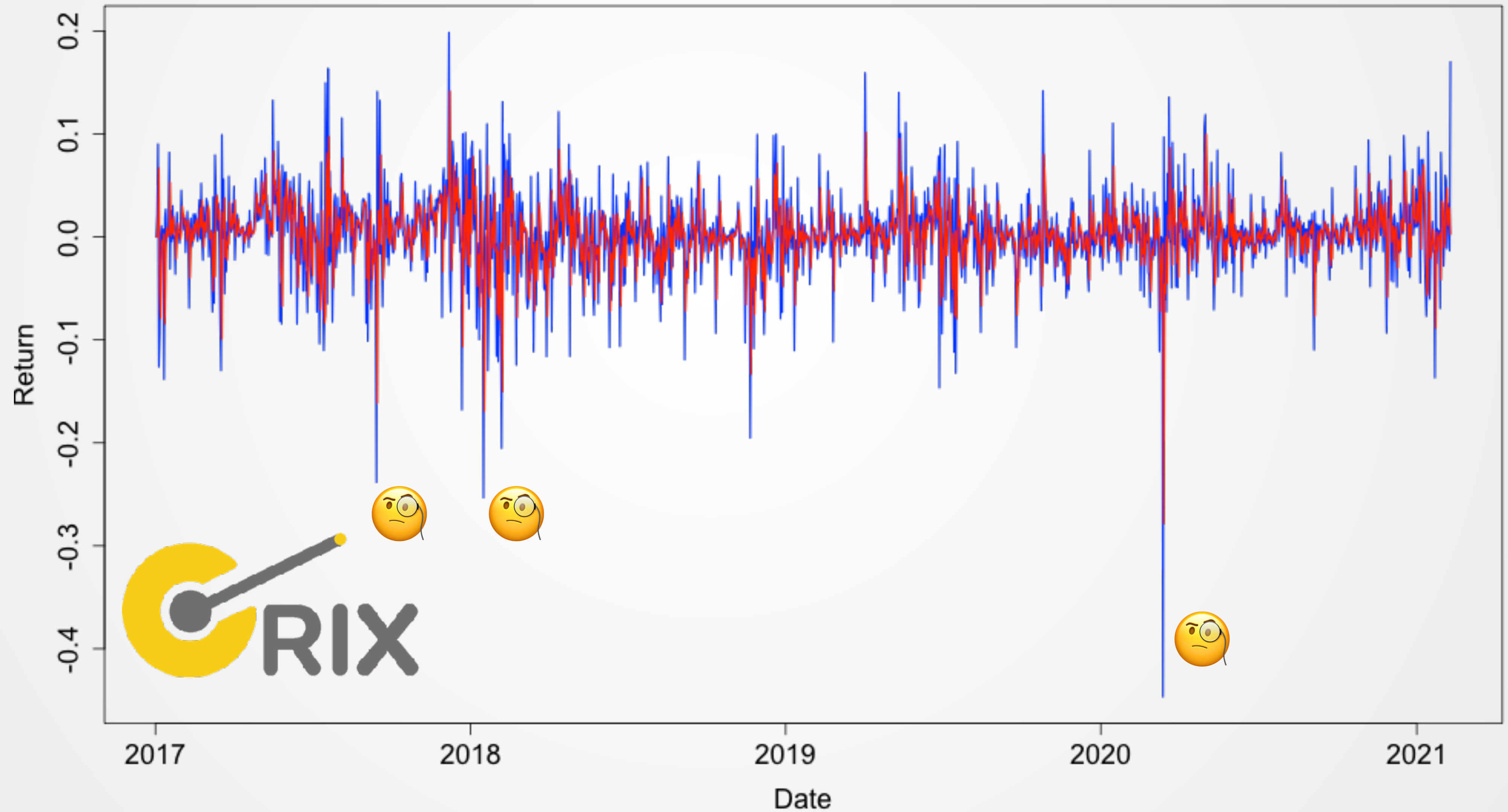
$$\hat{r}_{x,t+1}^- = \hat{r}_{x,t}$$

$$\hat{P}_{t+1}^- = \hat{P}_t + Q$$

$$\hat{r}_{y,t+1}^- = \hat{r}_{x,t+1}^-$$

Estimation Process KF_modelres

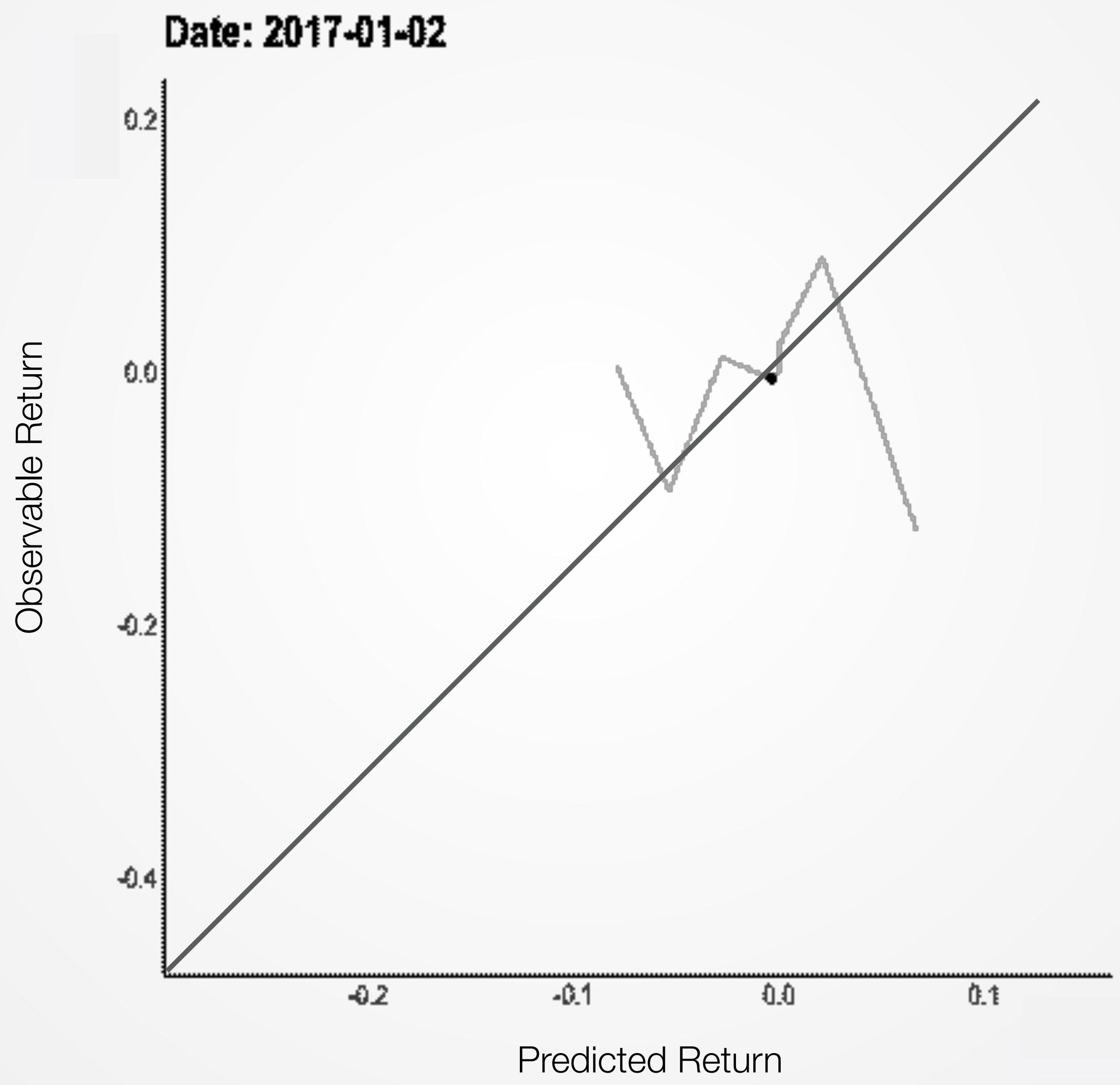
- $Q = 0.03 \;\; R = 0.03$



Return of CRIX: observable return, predicted return

Estimation Process KF_modelres

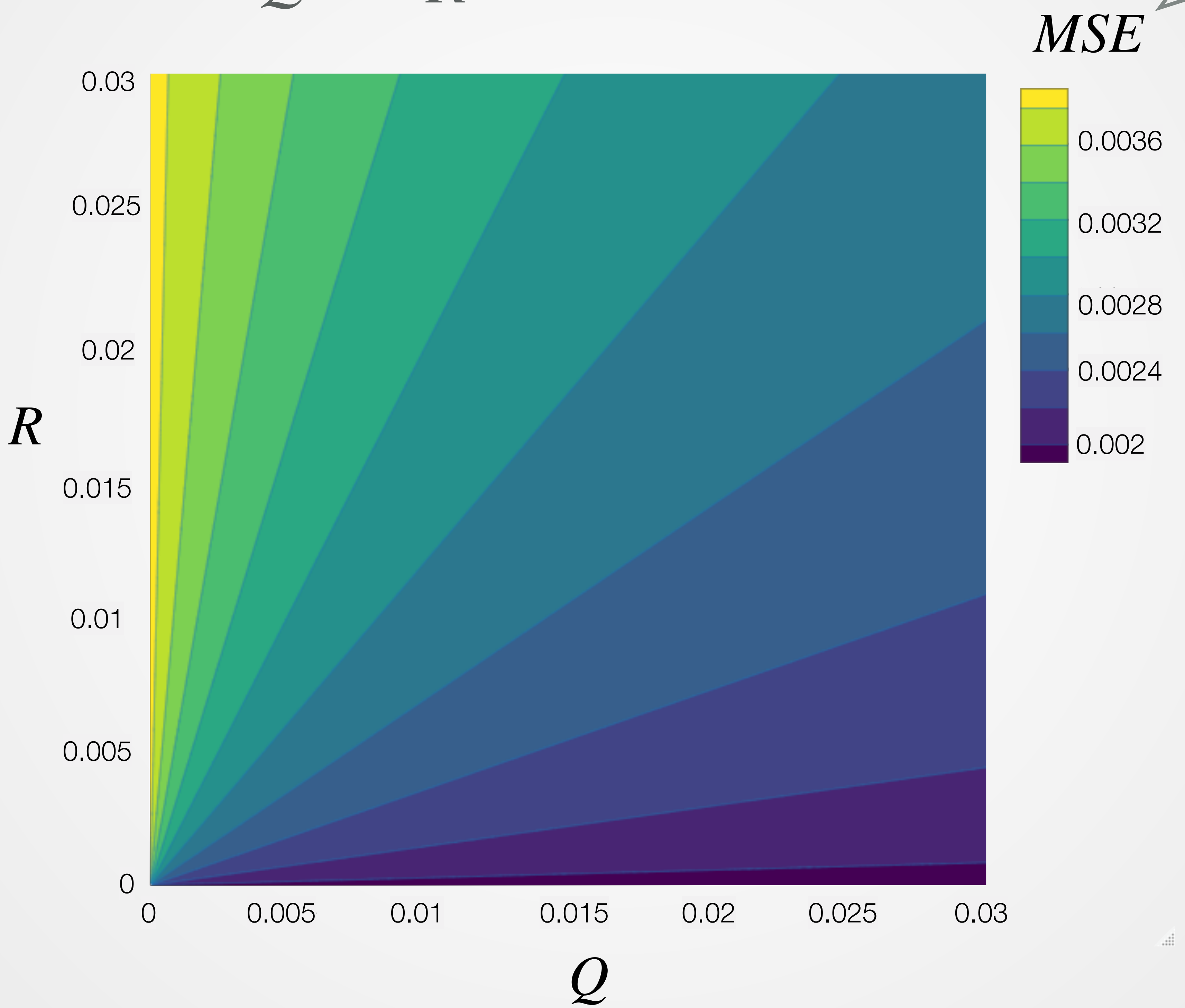
- $Q = 0.03 \;\; R = 0.03$



Estimation Process KF_modelres

$$\frac{1}{N}\sum_{t=1}^{N}(r_{y,t} - \hat{\bar{r}}_{x,t})^2$$

☐ MSE of Different Q and R



Other applications

- The common feature of inflation expectations (IE)

Bond Yield
- IE
  - UK
  - France
  - Germany
  - Italy
  - Sweden
- Risk Premiums

$\succ$ Common feature ?

Other applications

- The common factor of inflation expectations (IE) MTS

$$\Pi_{t,\tau} = q_\tau \Pi_{t-1,\tau} + p_\tau + v_{t,\tau}$$

$$\hat{\pi}_{it,\tau} = n_{i,\tau} \Pi_{t,\tau} + m_{i,\tau} + u_{it,\tau}$$

$\Pi_{t,\tau}$: Common feature of τ year IE at time t

$\hat{\pi}_{it,\tau}$: i country's τ year IE at time t

$m_{i,\tau}$, $n_{i,\tau}$, p_τ, and q_τ are unknown

$v_{it,\tau}$ and $u_{it,\tau}$ are i.i.d

Result of Extraction

3 year IE of five European countries, Common feature (Kalman Filter)

Reference

- Kalman RE (1960)
  A New Approach to Linear Filtering and Prediction Problems
  Journal of Basic Engineering, 82: 34–45.
- Kalman RE, Bucy RS (1961)
  New Results in Linear Filtering and Prediction Theory
  Journal of Basic Engineering, 83: 95–108.
- Nkomo R, Kabundi A (2013)
  Kalman Filtering and Online Learning Algorithms for Portfolio Selection
  Working paper 394, Economic Research Southern Africa

Reference

- Chen S, Härdle WK, Wang W (2021) The common and specific components of inflation expectations across European countries. Empir Econ (2021). https://doi.org/10.1007/s00181-021-02027-1
- https://www.zhihu.com/question/23971601
- https://www.bilibili.com/video/BV1ez4y1X7eR/?spm_id_from=333.788&vd_source=0f130fea32a288e2130afc13c0c051f8

Reference

1960	Filter in discrete time	Kalman, R. E.(1960). A new approach to linear filtering and prediction problems. *Journal of Basic Engineering Transactions, 82*, 35-45.
1961	Filter in continuous time	Kalman, R. E., & Bucy, R. S. (1961). New Results in Linear Filtering and Prediction Theory. *Trans. ASME, Ser. D, J. Basic Eng* (Vol.83, pp.109).
1960s	Use in air and space navigation (Schmidt- Kalman filter)	Schmidt, S. (1966). "Applications of State-space Methods to Navigation Problems". In Leondes, C. (ed.). *Advances in Control Systems*. **3**. New York, NY: Academic Press. pp. 293–340.
1960s	Use in non-linear field(Extended Kalman Filter，EKF)	Sunahara, Y. (1969). “An approximate method of state estimation for nonlinear dynamical systems,” in Proceedings of the Joint Automatic Control Conference, University of Colorado, Boulder, Colo, USA, Bucy, R. S. and Senne, K. D. (1971) “Digital synthesis of non-linear filters,” *Automatica*, vol. 7, no. 3, pp. 287–298, 1971
1989	Use in Neural Network	Singhal, S., & Wu, L. (1989). Training Multilayer Perceptrons with the Extended Kalman Algorithm. *neural information processing systems*.