# Change Point Detection with Switching Kalman Filter

Alptekin Orbay (in military service), Korhan Polat

## Problem & Motivation

Our aim is to identify trend changes in time series data. In other words, it can be seen as dividing a time series into meaningful segments. Simply, we can define the problem as

$$\textit{data} = \textit{model} + \textit{noise}$$

If we conceptualize the data as above, there are different ways that might lead to change point occurrences. Some of them are

- Order change
- Parameter Change
- Noise Level Change

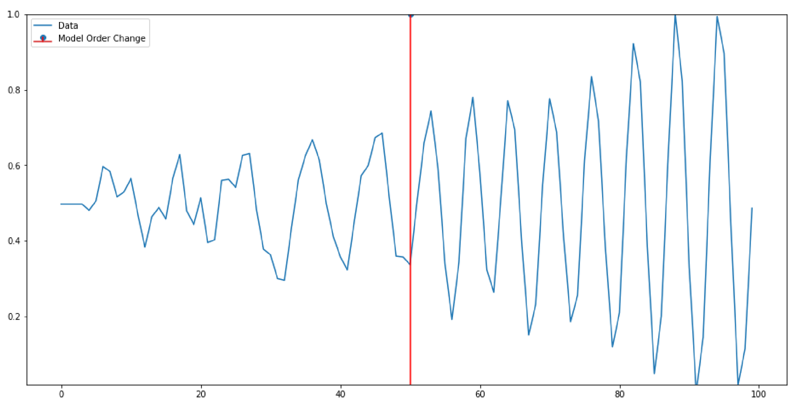

Figure 1: $4^{th}$ order AR model to $2^{nd}$ order AR model

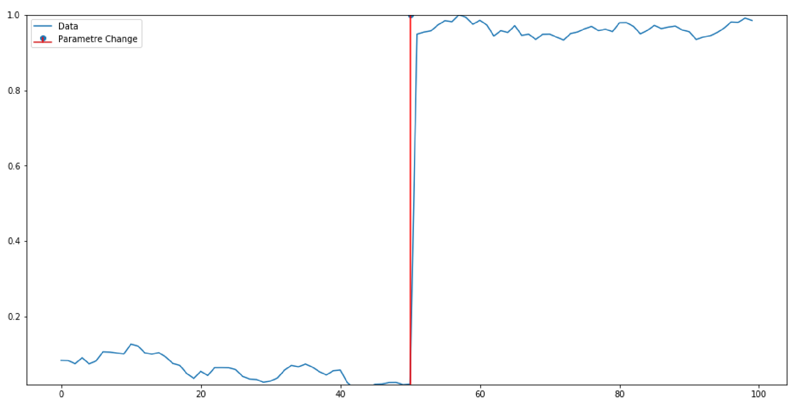

Figure 2: Change in the parameters of the distribution

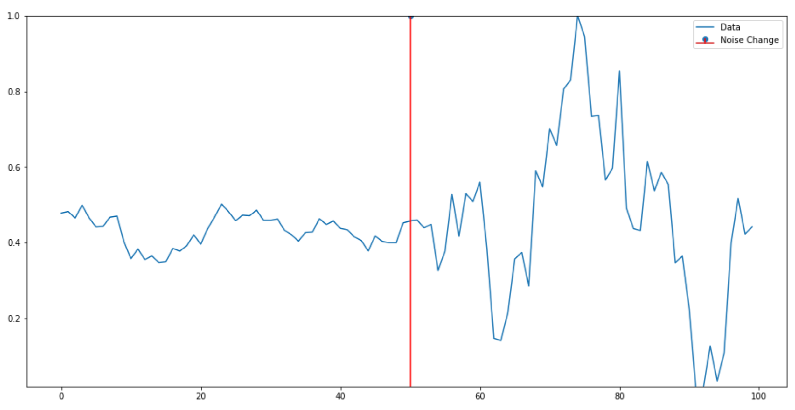

Figure 3: Noise level change

We will investigate how to detect change point in model parameters.

## Kalman Filter

- Prediction

$$x_{t|t-1} = Ax_{t-1|t-1}$$
$$P_{t|t-1} = AP_{t-1|t-1}A^\top + Q$$

- Update

$$S_t = CP_{t|t-1}C^\top + R$$
$$G_t = P_{t|t-1}C^\top S_t^{-1}$$
$$e_t = y_t - Cx_{t|t-1}$$
$$x_{t|t} = x_{t|t-1} + G_t e_t$$
$$P_{t|t} = P_{t|t-1} - G_t C P_{t|t-1}$$
$$l_t = l_{t-1} - 0.5\log|2\pi S_t| - 0.5 e_t^\top S_t^{-1} e_t$$

## References

[1] K. P. Murphy. Switching Kalman filters. Technical report, DEC/Compaq Cambridge Research Labs, 1998.

**github link:**
https://github.com/claudiusalp/SwitchingKalmanFilter

## Model

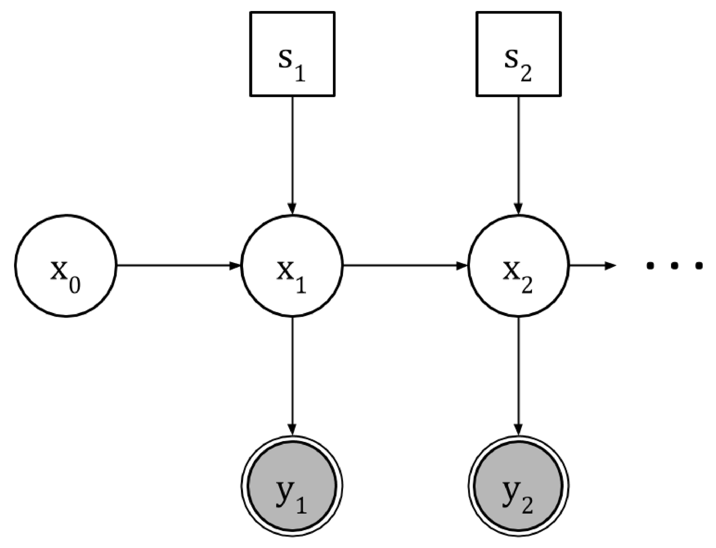


Figure 4: State transition model: square denotes discrete variables, circles for continuous variables and double circles are for observed variables.

$$x_t|x_{t-1}, s_t \sim \begin{cases} \mathcal{N}(x_t; Ax_{t-1}, Q), & s_t = 0 \\ \mathcal{N}(x_t; u_0, P_0), & s_t = 1 \end{cases}$$

$$y_t|x_t \sim \mathcal{N}(y_t; Cx_t, R)$$

- $A$ is transition matrix and $C$ is observation matrix.
- $Q$ is transition covariance matrix and $R$ is observation covariance matrix.
- $u_0$ is initial mean and $P_0$ initial covariance matrix.

## Method

The model tells that if $s_t = 0$, the state follows the preceding state, otherwise the state is switched to a new state from the initial distribution thus a new trend starts. Our solution is to adopt the Kalman Filter for this proposed model since we assume that hidden states and observations obey Gaussian distribution. There are two options for a state's next transition, to reset or not. For each time step, one of the hidden states is the reset state whereas the others are filtered normally. So, for time $t = t^*$, there are $t^* + 1$ possible hidden states. Each state has a likelihood coming from Kalman Filter with regard to previous state and a probability can be assigned to each one according to likelihoods.

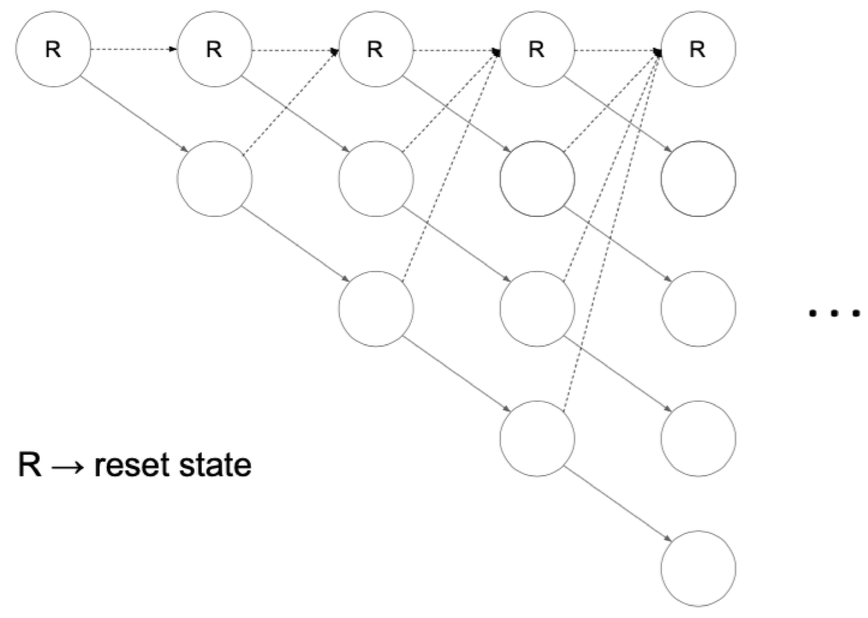


Figure 5: Possible hidden states and paths considering change point at each time step.

Note that, in our model the initial mean and covariance play important roles. The covariance should be flexible enough to allow changes. If it is too wide, estimated change points become noisy. If it is too sharp, change points become rare.

## Well Log Data

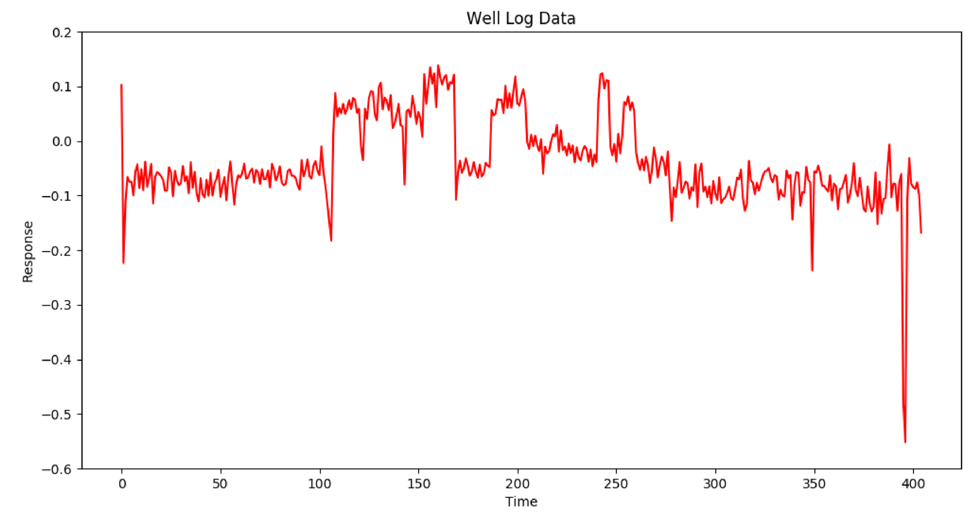


Figure 6: Well log data

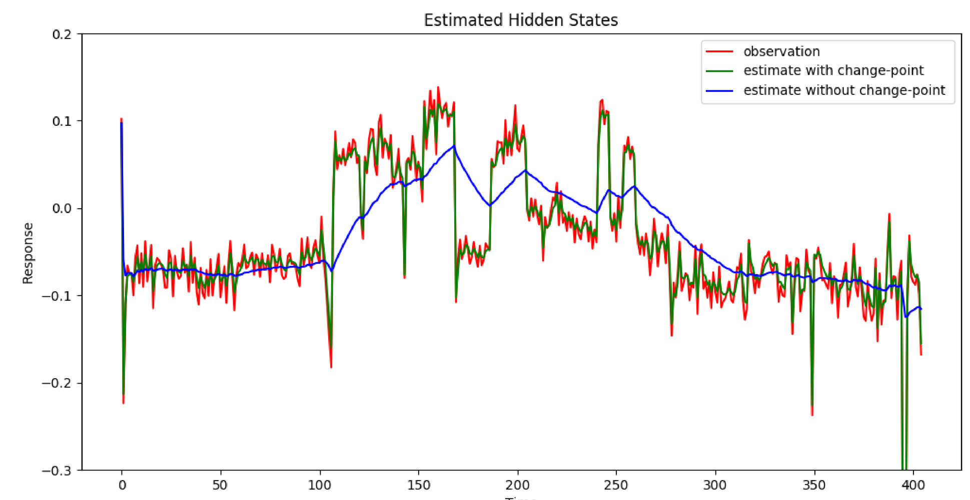


Figure 7: Estimated hidden states

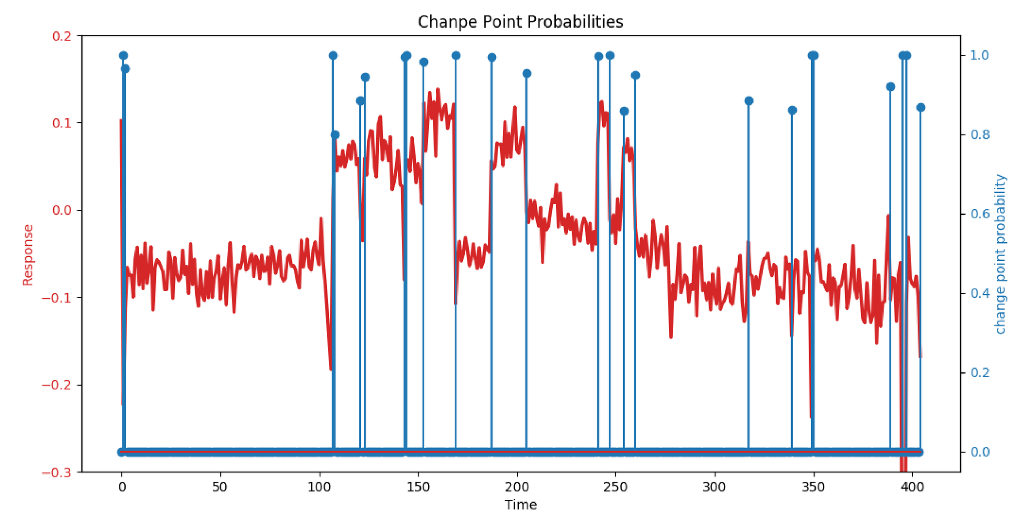

Figure 8: Change point probabilities

## Bitcoin

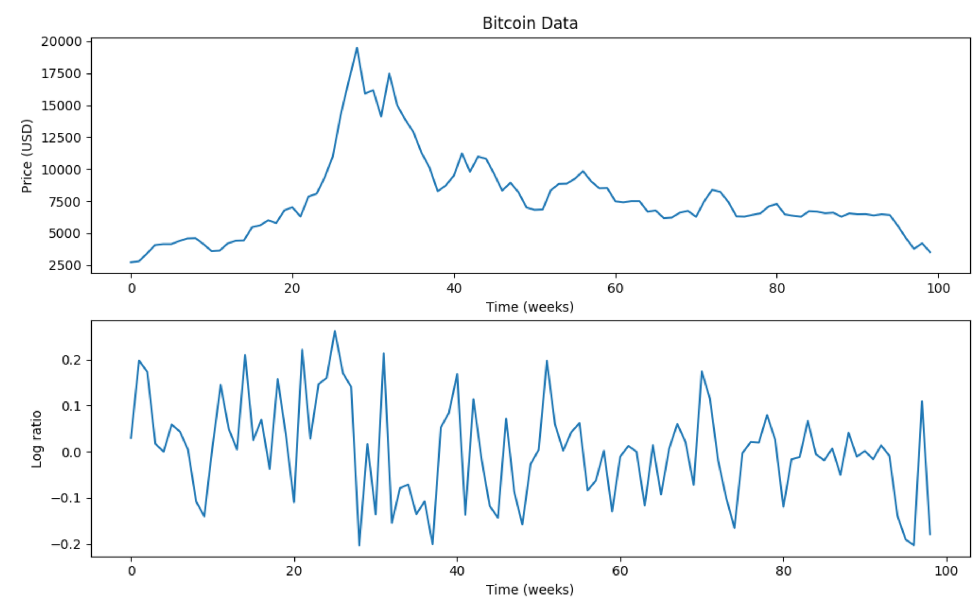


Figure 9: Historical BTC data since August 2017

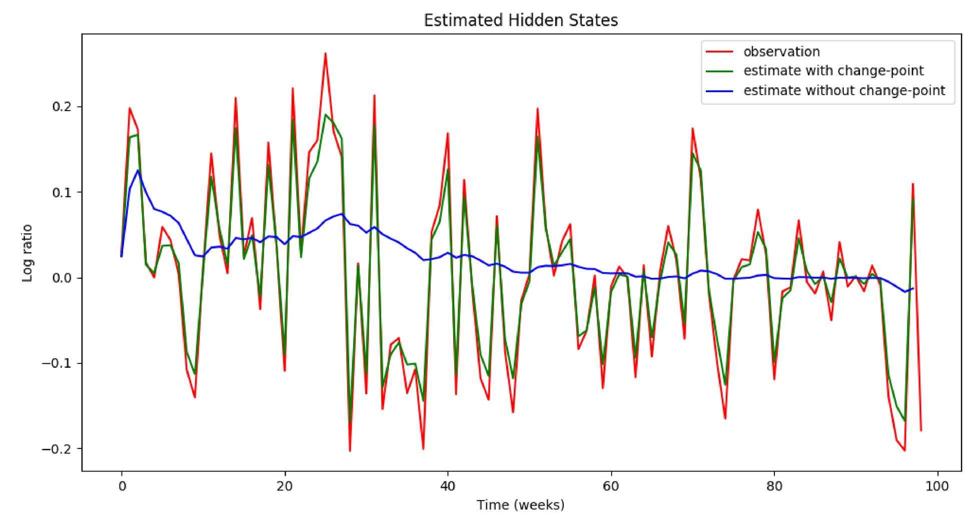


Figure 10: Estimated hidden states

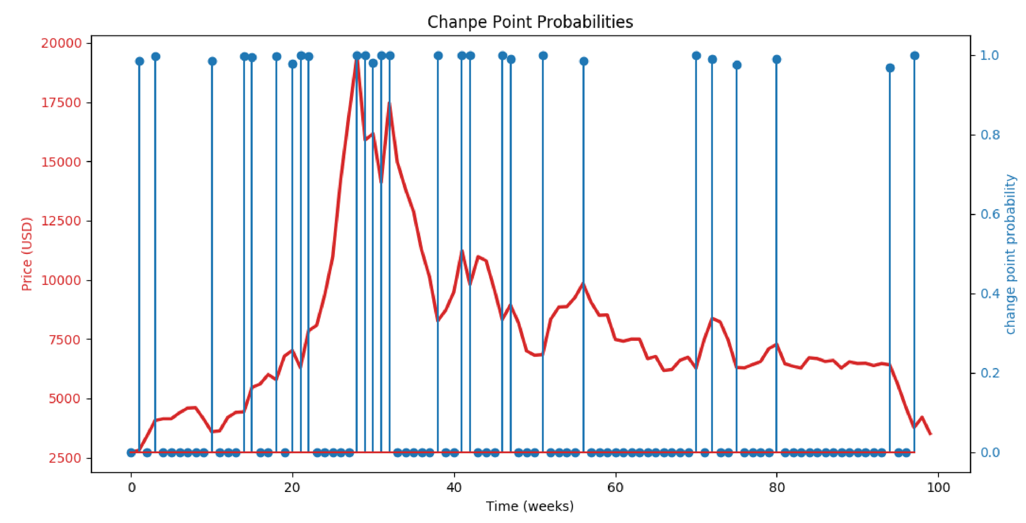


Figure 11: Change point probabilities