Kalman Filter

Kalman Filter Fair Price Estimation

- The observed mid-price is an important reference for fair price, but they tend to be noisy, dynamic and unstable, especially for a low liquidity market.
- The Kalman filter is an algorithm used for estimating the state of a dynamic system from a series of incomplete and noisy measurements.

State Transition Equation

$$\mathbf{x}(t) = \mathbf{F}(t)\mathbf{x}(t\text{-}1) + \mathbf{w}(t)$$

- x(): true state
- F(): state-transition matrix
- $\mathbf{w}(t) \sim_{iid} N(0, Q(t))$ process noise

Let x() represents the fair price, F() = 1:

$$\mathbf{x}(t) = \mathbf{x}(t\text{-}1) + \mathbf{w}(t)$$

- fair price follows random walk

Measurement (Observation) Equation

$z(t) = H(t)x(t) + v(t)$

- x(): true state
- z(): observation of the true state (noisy)
- H(): observation matrix, (true state space -> observed space)
- $v \sim_{iid} N(0, R(t))$ observation noise

Let z(t) is the observed price, H() = 1:

$z(t) = x(t) + v(t)$

- The observed price is the true price plus some observation noise

State Update Equation

$$x(t|t) = x(t|t\text{-}1) + K(t)\,(z(t) - x(t|t\text{-}1))$$
$$= (1\text{-}K(t))\,x(t|t\text{-}1) + K(t)\,z(t)$$

- $x(t|t)$: posterior state estimate
- $x(t|t\text{-}1)$: prior state estimate = $x(t\text{-}1|t\text{-}1)$
- $z()$: observation of the true state
- $z(t) - x(t|t\text{-}1))$: innovation
- $K()$: Kalman gain

This is a recursive estimator, no history of observations or estimates is required.

This is a weighted average between the prior estimate and the new observation, with more weight being given to estimates with greater certainty. The weight is controlled by Kalman Gain.

Optimal Kalman Gain K(t)

$$K(t) = P(t|t\text{-}1) \,/\, [P(t|t\text{-}1) + R(t)]$$
$$P(t|t\text{-}1) = P(t\text{-}1|t\text{-}1) + Q(t)$$
$$P(t|t) = [1\text{-}K(t)]\, P(t|t\text{-}1)$$

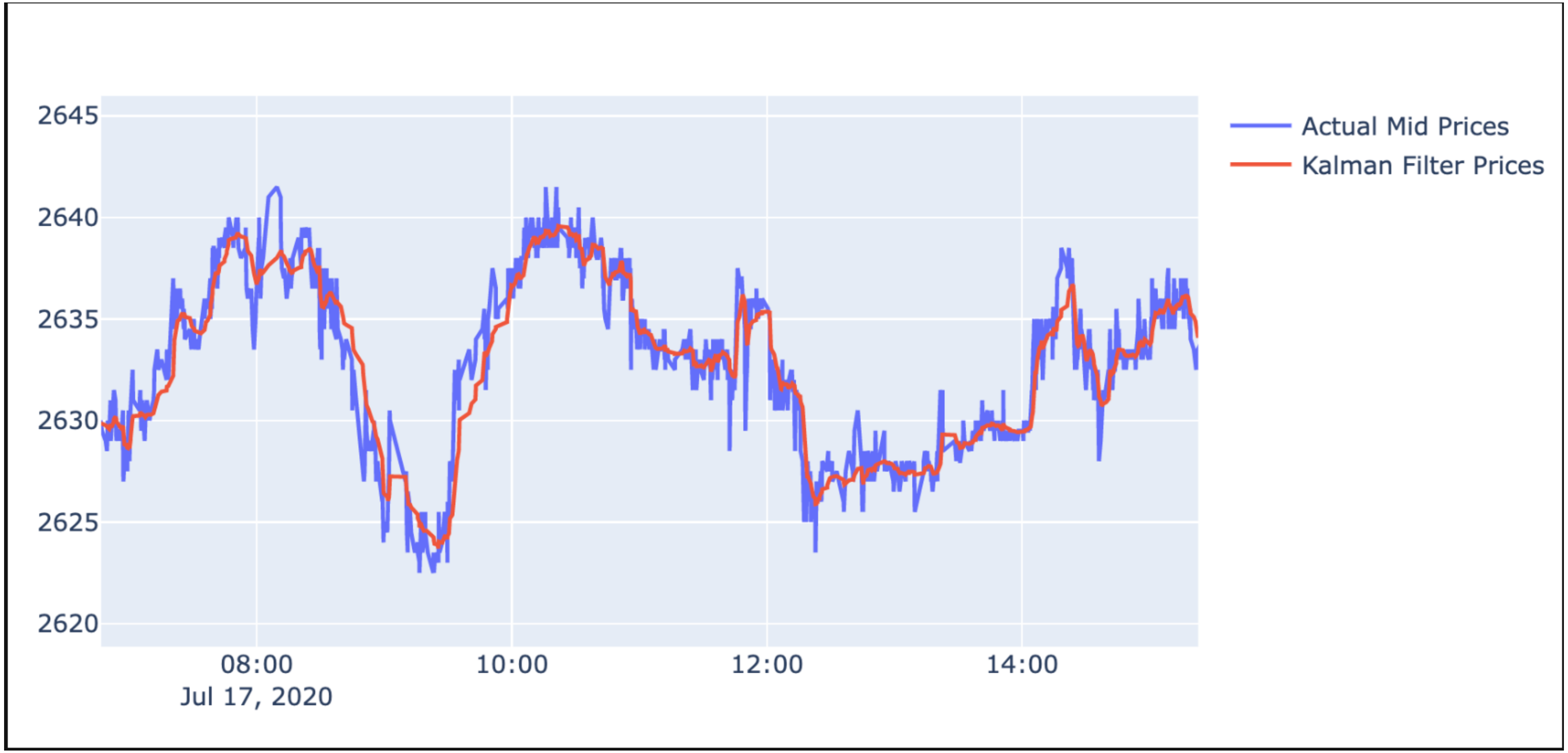
- $P(t|t\text{-}1)$: prior covariance (estimated accuracy of the state estimate)
- $P(t|t)$: updated posterior estimate covariance
- $[P(t|t\text{-}1) + R(t)]$: innovation covariance

Kalman Gain determines how much we trust the new observation

With a high gain, the filter places more weight on the most recent measurements. With a low gain, the filter conforms to the model predictions more closely

Kalman Gain may need to be tuned to achieve a particular performance

Actual Mid Prices VS Kalman Filter Prices



Avellaneda and Stoikov’s Market Making Strategy

Avellaneda and Stoikov’s Strategy

- ["High-frequency trading in a limit order book"]
- Avellaneda & Stoikov focus on market maker's two challenges:
  - inventory risk
  - optimal bid and ask spreads
- Algorithm:
  1. Calculate a reservation price based on inventory, volatility, time etc.
  2. Calculate the optimal bid and ask spread based on utility framework and the microstructure of actual limit order books
  3. Create market orders using the reservation price as reference:
     - Bid offer price = reservation price - optimal spread / 2
     - Ask offer price = reservation price + optimal spread / 2

Reservation price

$$r(s, q, t) = s - q\gamma\sigma^2(T - t)$$

- s = current market mid price
- q = quantity of assets in inventory of base asset (could be positive/negative for long/short positions)
- σ = market volatility
- T = closing time
- t = current time
- γ = inventory risk aversion parameter

Creating symmetrical bid/ask orders around the market mid-price may lead to inventory skewing in one direction

Instead of using market mid-price, we can calculate a new reference price that takes inventory, market volatility, and time until trading session ends into consideration

Optimal Spread

$$\delta^a + \delta^b = \gamma\sigma^2(T - t) + \frac{2}{\gamma}ln(1 + \frac{\gamma}{\kappa})$$

- δa, δb = bid/ask spread, symmetrical → δa=δb
- γ = inventory risk aversion parameter
- σ = market volatility
- κ = order book liquidity parameter

Denser order book **->** significant κ value **->** smaller optimal spread (more competition)