# Initial Guess

Renzo Miguel Caballero Rosas

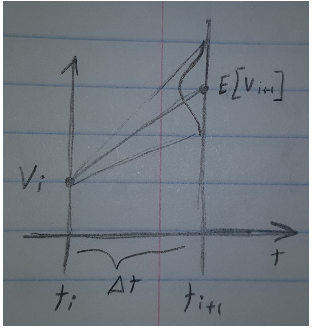
March 14, 2020

# Quadratic mean minimization

We consider the transition $\Delta V_i = V_{i+1} - V_i$ with $\Delta t = t_{i+1} - t_i$. $(V_{i+1}|V_i)$ is a random variable which conditional mean can be approximated by the solution of the system

$$\begin{cases} \mathrm{d}\mathbb{E}[V] & = -\theta_t \mathbb{E}[V]\mathrm{d}t \\ \mathbb{E}\left[V(t_i)\right] & = V_i, \end{cases}$$

evaluated in $t_{i+1}$ (i.e., $\mathbb{E}\left[V(t_{i+1})\right]$). Then, the random variable $(V_{i+1} - \mathbb{E}\left[V(t_{i+1})\right])$ has approximately zero mean.




If we assume that $\theta_t = \theta_0$ for all $t \in [t_i, t_{i+1}]$, then $\mathbb{E}\left[V(t_{i+1})\right] = V_i e^{-\theta_0 \Delta t}$. If we have a total of $n$ transitions, we can write the regression problem for the conditional mean with $L^2$ loss function as

$$\theta_0^* \approx \arg\min_{\theta_0} \left[ \sum_{i=1}^{n} \left( V_{i+1} - \mathbb{E}\left[V(t_{i+1})\right] \right)^2 \right] = \arg\min_{\theta_0} \left[ \sum_{i=1}^{n} \left( V_{i+1} - V_i e^{-\theta_0 \Delta t} \right)^2 \right]. \quad (1)$$

# Quadratic mean minimization

We take the first order approximation w.r.t. $\theta_0$

$$e^{-\theta_0 \Delta t} = 1 - \theta_0 \Delta t + \mathscr{O}(\theta_0^2),$$

and introduce it in equation (1). We get

$$\theta_0^* \approx \arg\min_{\theta_0} \underbrace{\left[\sum_{i=1}^{n} \left(V_{i+1} - V_i(1 - \theta_0 \Delta t)\right)^2\right]}_{:= f(\theta_0)}. \tag{2}$$

As $f(\theta_0)$ is convex in $\theta_0$, solving (2) (finding $\theta_0^*$) is equivalent to solving $\frac{\partial f}{\partial \theta_0}(\theta_0^*) = 0$.

# Quadratic mean minimization

$$
\begin{aligned}
\frac{\partial f}{\partial \theta_0} &= \sum_{i=1}^{n} 2(-V_i)(-\Delta t)(V_{i+1} - V_i(1 - \theta_0 \Delta t)) \\
&= \sum_{i=1}^{n} 2 V_i \Delta t (V_{i+1} - V_i(1 - \theta_0 \Delta t) \\
&= \sum_{i=1}^{n} 2 V_{i+1} V_i \Delta t - 2 V_i^2 \Delta t + 2 V_i^2 \Delta t^2 \theta_0 .
\end{aligned}
$$

Then, $\theta_0^*$ satisfies

$$
\theta_0^* \approx \frac{\sum_{i=1}^{n} V_i (V_i - V_{i+1})}{\Delta t \cdot \sum_{i=1}^{n} (V_i)^2} . \tag{3}
$$

Notice that $\theta_0^*$ has dimension $\textbf{time}^{-1}$.

# Quadratic variation

One more time, we approximate the SDE by its E-M scheme. In particular, we approximate the Itô quadratic variation with the discrete one:

- Itô process quadratic variation: $[V]_t = \int_0^t \sigma_s^2 \, \mathrm{d}s$.
- Discrete process quadratic variation: $[V]_t = \sum_{0<s\leq t} (\Delta V_s)^2$.

Then, considering $\Delta t$ the time between measurements, we approximate:

$$\theta_0^* \alpha^* \approx \frac{\sum_{i=1}^n (\Delta V_i)^2}{2\Delta t \sum_{i=1}^n (V_i + p_i)(1 - V_i - p_i)}. \tag{4}$$

# Initial parameters

We first approximate $\theta_0^* \alpha^*$ from (4). Then, using all the data, assuming $\theta_t = \theta_0$ and equation (3), we can approximate $\theta_0^*$. However, as we have $\theta_0^* \alpha^*$, we can check for which days the assumption $\theta_t = \theta_0$ was wrong, and remove that days from all the data.

We repeat the process until the assumption is correct.

# Boundedness of $\theta_t$

To guarantee an unique solution for the process $X_t$, we need boundedness in $\theta_t$ for $t \in [0, T]$. For our definition

$$\theta_t = \max\left(\theta_0, \frac{\alpha\theta_0 + |2\dot{p}_t|}{2\min(1-p_t, p_t)}\right), \tag{5}$$

we do not have a bound for $\theta_t$ if $p_t \to 0^+$ or $p_t \to 1^-$. However, if we ensure that $p_t \in [\delta, 1-\delta]$ for some $0 < \delta < \frac{1}{2}$ and all $t \in [0, T]$, then $\theta_t < M(\delta) < \infty$ for all $t \in [0, T]$. So, we define the corrected forecast

$$p_t^\delta = \begin{cases} \delta & \text{if} \quad p_t < \delta \\ p_t & \text{if} \quad \delta \le p_t < 1-\delta \\ 1-\delta & \text{if} \quad p_t \ge 1-\delta, \end{cases} \tag{6}$$

where we choose and fix $\delta = 10^{-2}$.

# Boundedness of $\theta_t$

If we use the corrected forecast (6) in our definition (5), we can achieve a new (bounded in $[0, T]$) definition

$$\theta_t^{\delta} = \max\left(\theta_0, \frac{\alpha\theta_0 + |2\dot{p}_t^{\delta}|}{2\min(1 - p_t^{\delta}, p_t^{\delta})}\right). \tag{7}$$

Using our new definition $\theta_t^{\delta}$ in the drift of our SDE:

- We guarantee existence and uniqueness of the solution, and that $X_t \in ]0,1[$ for all $t \in [0,1]$.
- As $X_t \in ]0,1[$, the Lamperti transform is well defined for all $t \in [0, T]$ and we can use Itô's lemma to compute the SDE for the transformed process.

# Estimation of $\delta$

Recall that to estimate $\theta_0$ in slide (2), we assume that $\theta_t = \theta_0$. From equation (7) we can notice that the assumption depends on $\theta_0$, $\alpha$, and $\delta$.

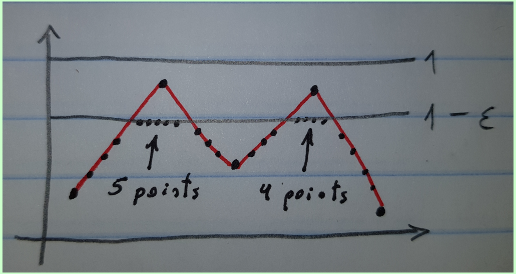
From the quadratic variation (slide 5, equation 4), we can estimate, using all the data, the product $\theta_0^*\alpha^*$. We call it $k = \theta_0^*\alpha^*$. The importance of this value is that we only compute it once and using all the data.

Now, with $k$ and some initial $\delta^{(0)}$, we choose only the data where the assumption $\theta_t = \theta_0$ holds, and we estimate $\theta_0^*$ using the quadratic mean minimization (slide 4, equation 3). Then, we choose part of the data where the assumption was not true. If also, we have that $|\dot{p}_t^\delta| = 0$, then equation 3 gives us $\theta_t^{\delta *} \in \mathbb{R}^+$. As we were assuming that $\theta_t^{\delta *} \neq \theta_0$, then the next equality should hold

$$\theta_t^{\delta *} = \frac{\theta_0^*\alpha^*}{\delta^{(0)}} = \frac{k}{\delta^{(0)}}. \quad \text{Then, we can estimate} \quad \delta^{(1)} = \frac{k}{\theta_t^{\delta *}} \quad \text{and repeat all until} \quad \delta^{(i+1)} \approx \delta^{(i)}.$$

# Estimation of $\delta$

To guarantee that $|\dot{p}^{\delta}_{t_i}| = 0$, we need that $p_{t_i} = p_{t_{i+1}}$. Then, if we have $n+1$ truncated consecutive forecasts (i.e., $p_{t_i}, \ldots, p_{t_{i+n}}$), we can use the $n$ transitions $\Delta V_{t_i}$ because $|\dot{p}^{\delta}_{t_i}| = \cdots = |\dot{p}^{\delta}_{t_{i+n-1}}| = 0$, but $|\dot{p}^{\delta}_{t_{i+n}}| \neq 0$.




It may happen that $\delta \to 0$ because as $\delta$ becomes smaller, the forecasts $p_t \notin [\delta, 1-\delta]$ leads to larger $\theta_t$ because $\theta_t \approx \frac{\theta_0 \alpha}{\delta}$.

# Estimation of $\delta$

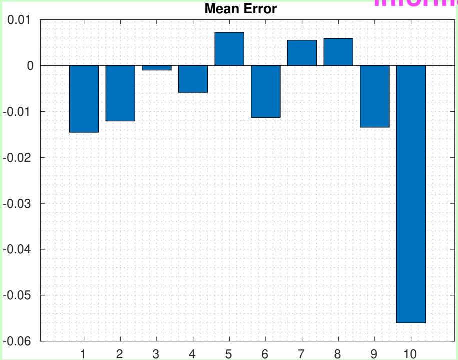


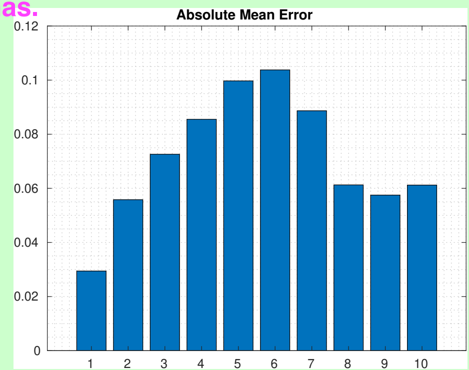
esto es interesante, tendremos que pensar mas al respecto para ver si todo es intuitivo

We start with $\delta^{(0)} = 0.1$, and finish with $\delta^{(10)} = 6.07 \times 10^{-3}$. Also, $\max_{[0,1]} \theta_t^{\delta^{(10)}} \approx 16$. For $\delta < 6.12 \times 10^{-3} = \delta^*$ we do not have to truncate the forecast for any value during 2019. The iteration stopped because $\delta^{(10)} < \delta^*$, and $p_t \in [\delta^{(10)}, 1 - \delta^{(10)}]$ for all forecasts.

# Extra: Interesting data processing

**por favor agrega un scatter plot de forecast error versus p_t, sin promedios, eso informa mas.**




What we are seeing is the **mean error** and **mean absolute error** as a function of the forecast. This is, for each interval with length 0.1 (i.e., [0,0.1), [0.1,0.2), etc.), we average all the errors corresponding to measurement where the forecast was in that intervals, and after we average over the number of elements in each interval. In some future, we can use this information to construct an even more realistic model.