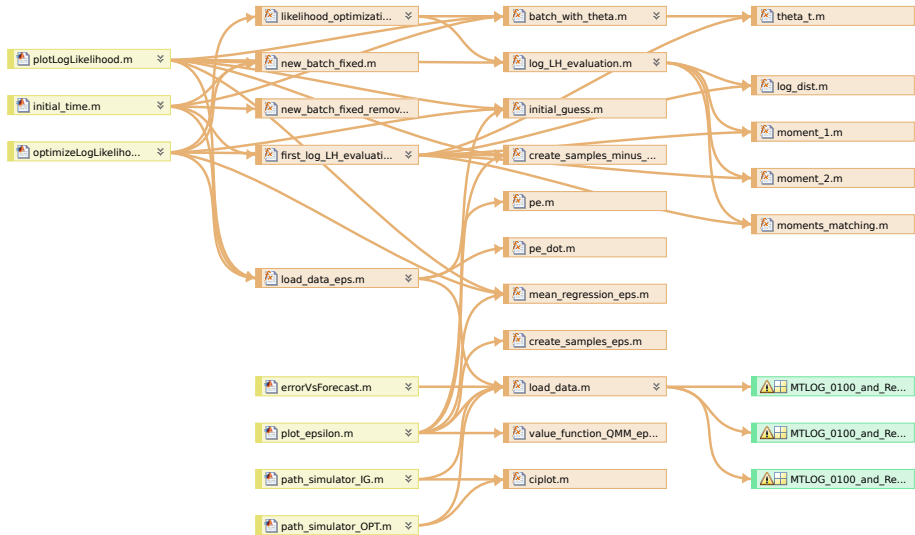


# MATLAB: Read Me

Renzo Miguel Caballero Rosas

August 4, 2020

# Code diagram:



## Questions:

- ▶ Why  $dX = a(X; \theta)dt + b(X; \theta, \alpha)dW$  and not  $dX = a(X; \theta)dt + b(X; \gamma)dW$ ? Where all  $\theta, \alpha, \gamma \in \mathbb{R}^+$ . Because in the way it is defined,  $\theta$  controls the mean reversion and  $\alpha$  the width of the confidence band. However, maybe it is better to optimize over  $\theta$  and  $\gamma$ ? Because of the relative dimension. After, trivially we can compute  $\alpha = \gamma/\theta$ .
- ▶ Which is Beta, the measurements or the transitions?
- ▶ Which data in the histograms? Measurements or transitions?
- ▶ I had to add an artificial correction to ensure that the second moment is always positive. It is in function log\_LH\_evaluation.m.

Some keywords:

- ▶ Our process is: **High-frequency in a fixed time-interval.**

## Normalization:

Given the SDE

$$dV_t = -\theta_t V_t dt + \sqrt{2\theta_0 \alpha(V_t + p_t)(1 - V_t - p_t)} dW_t,$$

we consider the normalized differentials  $d\hat{t} = \frac{dt}{T}$ , and  $d\hat{W}_t = \frac{dW_t}{\sqrt{T}}$ . Then, we can write the SDE as

$$dV_t = -\theta_t T V_t d\hat{t} + \sqrt{2\theta_0 T \alpha(V_t + p_t)(1 - V_t - p_t)} d\hat{W}_t.$$

We conclude that, whatever the normalization constant  $T$  is, it gets absorbed by the parameter  $\theta_t$  or  $\theta_0$  (let  $\hat{\theta}_t = \theta_t T$  and  $\hat{\theta}_0 = \theta_0 T$ ).

The E-M representation is

$$V_{t_{n+1}} = V_{t_n} - [\hat{\theta}_{t_n} V_{t_n}] \Delta s + \left[ \sqrt{2\hat{\theta}_0 \alpha(V_{t_n} + p_{t_n})(1 - V_{t_n} - p_{t_n})} \right] \sqrt{\Delta s} \Delta \hat{W}_{t_n}, \quad V_{t_0} = v_0,$$

for  $n \in \{0, \dots, m-1\}$ ,  $t_j = t_0 + j\Delta s$ ,  $t_0 = 0$ ,  $t_m = 1$ , and  $\Delta \hat{W}_{t_n}$  normal (0,1) for each  $t_n$ .

## SDE first moment (1/2):

Given some measurement  $v_{t_{n-1}}$ , we want to compute the first moment at time  $t_n$ . The exact first moment  $m_1(s)$  for  $s \in [t_{n-1}, t_n]$  is the solution of the ODE

$$\begin{cases} dm_1(s) = [-m_1(s)\theta(s)] ds, \\ m_1(t_{n-1}) = v_{t_{n-1}}. \end{cases}$$

If  $\theta(t_{n-1}) = \theta(t_n) = \theta$ , the solution is  $m_1(t_n) = m_1(t_{n-1})e^{-\theta(t_n - t_{n-1})}$ .

Otherwise, we compute a linear interpolation for  $\theta(s)$  and solve the ODE using Forward-Euler:

$$m_1(s_n) = m_1(s_{n-1})(1 - \theta(s_{n-1})\Delta s).$$

## SDE first moment (2/2): CODE

```
1 function m1 = moment_1(v, theta_0, alpha, th1, th2, p1, p2, dt, n)
2     % 02/04/2020 15:16
3
4     if th1 == th2 % We have the exact solution.
5         m1 = v*exp(-th1*dt);
6     else % Otherwise, we compute F-E.
7         p = @(i) p1 + (p2-p1) * i/n; % We interpolate p(t).
8         pe_dot = (p2-p1) / dt;
9         theta = @(i) theta_t(theta_0, alpha, p(i), pe_dot);
10        m1(1) = v;
11
12        ds = dt/n;
13        for i = 2:n
14            m1(i) = m1(i-1) * (1 - theta(i-1)*ds);
15        end
16    end
17
18 end
```

## SDE second moment (1/2):

Given some measurement  $v_{t_{n-1}}$ , we want to compute the second moment at time  $t_n$ . The exact second moment  $m_2(s)$  for  $s \in [t_{n-1}, t_n]$  is the solution of the ODE

$$\begin{cases} dm_2(s) &= [-2m_2(s)(\theta(s) + \alpha\theta_0) + 2\alpha\theta_0m_1(s)(1 - 2p(s)) + 2\alpha\theta_0p(s)(1 - p(s))] ds, \\ m_2(t_{n-1}) &= v_{t_{n-1}}^2. \end{cases}$$

We compute a linear interpolation for the functions  $\theta(s)$  and  $p(s)$ . After, we solve the ODE using Forward-Euler:

$$m_2(s_n) = m_2(s_{n-1}) + [-2m_2(s_{n-1})(\theta(s_{n-1}) + \alpha\theta_0) + 2\alpha\theta_0m_1(s_{n-1})(1 - 2p(s_{n-1})) + 2\alpha\theta_0p(s_{n-1})(1 - p(s_{n-1}))] \Delta s.$$

We use the same discretization points for both  $m_1(s)$  and  $m_2(s)$ .



## SDE second moment (2/2): CODE

```
1 function m2 = moment_2(v, theta_0, th1, th2, p1, p2, alpha, m1, dt, n) % 02/04/2020 15:13
2
3     p      = @(i) p1 + (p2-p1) * i/n; % We interpolate p(t).
4     pe_dot = (p2-p1) / dt;
5     m2(1)  = v^2;
6     ds     = dt/n;
7
8     if th1 == th2 % In this case, m_1 has exact solution and theta_t = theta_0.
9         theta = @(i) th2;
10        m1     = @(i) v*exp(-th1*dt*(i/n));
11    else
12        theta = @(i) theta_t(theta_0, alpha, p(i), pe_dot);
13    end
14
15    for i = 2:n
16        m2(i) = m2(i-1) + (-2*m2(i-1)*(theta(i-1)+alpha*theta_0) + ...
17            2*alpha*theta_0*m1(i-1)*(1-2*p(i-1)) + ...
18            2*alpha*theta_0*p(i-1)*(1-p(i-1))) * ds;
19    end
20
21 end
```

## Density next measurement (1/2):

We want the next measurement  $V_{t_n}|V_{t_{n-1}}$  to have a Beta distribution, but with support in  $[a, b] = [-1, 1]$ . Given  $X \sim \beta(\xi_1, \xi_2)$  a **Beta distributed random variable**, we define the new random variable  $V = a + (b - a)X$  with support in  $[-1, 1]$ , and PDF  $f_V(v)$ .

We can compute:

$$\blacktriangleright \mathbb{E}[V] = a + (b - a)\mathbb{E}[X] = a + (b - a)\frac{\xi_1}{\xi_1 + \xi_2} = \mu_V.$$

$$\blacktriangleright \mathbb{V}[V] = (b - a)^2\mathbb{V}[X] = \frac{(b - a)^2\xi_1\xi_2}{(\xi_1 + \xi_2)^2(\xi_1 + \xi_2 + 1)} = \sigma_V^2.$$

Then, we want the SDE and our new PDF  $f_V(v)$  to have the same moments at each  $t \in \{\text{some appropriate domain}\}$ , i.e.,  $\mu(t) = m_1(t)$  and  $\sigma^2(t) = m_2(t) - m_1^2(t)$ .  $\mu(t)$  and  $\sigma^2(t)$  refers to the mean and variance of  $V_{t_n}|V_{t_{n-1}}$ , following the structure described for  $f_V(v)$ .

## Density next measurement (2/2):

For each measurement  $V_{t_{n-1}}$ , we can find the analytical moments for the SDE at time  $t_n$  solving the ODEs from slides 6 and 8. Then, we can find the parameters  $\xi_1$  and  $\xi_2$  such that both the SDE and the PDF of  $V_{t_n}|V_{t_{n-1}}$  have the same first and second moments at time  $t_n$ .

- ▶  $\xi_1 = -\frac{(1+\mu)(\mu^2+\sigma^2-1)}{2\sigma^2}$  all evaluated at time  $t_n$  (verified in **Mathematica 11.0**<sup>1</sup>).
- ▶  $\xi_2 = \frac{(\mu-1)(\mu^2+\sigma^2-1)}{2\sigma^2}$  all evaluated at time  $t_n$  (verified in **Mathematica 11.0**).

```
1 function [xi1,xi2] = moments_matching(m1,m2) % 02/02/2020 19:17
2 %      disp(['m1 = ',num2str(m1),' and m2 = ',num2str(m2), '.']);
3     mu    = m1;
4     sig2  = m2 - m1^2;
5     xi1   = - ((mu+1)*(mu^2+sig2-1)) / (2*sig2);
6     xi2   =  ((mu-1)*(mu^2+sig2-1)) / (2*sig2);
7
8 end
```

<sup>1</sup>File: [matchingVerification.nb](#).

## Log-density (1/2):

Recall the PDF  $f_V(v)$  from slide 10. We will use this density to model the random variables  $V_{t_n}|V_{t_{n-1}}$ . For  $[a, b] = [-1, 1]$ , we have that

$$f_V(v) = f_X(g^{-1}(v)) \left| \frac{d}{dv} g^{-1}(v) \right| \quad \text{where} \quad f_X(x) = \text{Beta}(\xi_1, \xi_2) \quad \text{and} \quad g(x) = a + (b-a)x.$$

$$\text{Then, } f_V(v) = \frac{1}{|(b-a)|} \frac{1}{B(\xi_1, \xi_2)} \left( \frac{v-a}{b-a} \right)^{\xi_1-1} \left( 1 - \frac{v-a}{b-a} \right)^{\xi_2-1} \quad \text{because } g^{-1}(v) = \frac{v-a}{b-a}.$$

Also, we have that (up to some constant values)

$$\log(f_V(v)) = \log\left(\frac{1}{B(\xi_1, \xi_2)}\right) + (\xi_1 - 1) \log\left(\frac{v-a}{b-a}\right) + (\xi_2 - 1) \log\left(\frac{b-v}{b-a}\right),$$

where  $\xi_1$  and  $\xi_2$  depends on the SDE moments.

## Log-density (2/2): CODE

```
1 function [val] = log_dist(v,xi1,xi2) % 03/02/2020 11:20
2     a    = -1;
3     b    = 1;
4     val = -betaln(xi1,xi2) + (xi1-1)*log((v-a)/(b-a)) + ...
5         (xi2-1)*log((b-v)/(b-a));
6 end
```

Notice we use the function **betaln(a,b)**. It is important to compute the log directly for the beta function, and not first the beta, and after apply log.

## Log-likelihood (1/2):

We introduce the number of paths  $M$ , and the number of measurements per path  $N+1$  ( $N$  transitions). Then, we have a total of  $M \times N$  samples to use. Notice that each pair  $(\xi_1, \xi_2)$  depends on  $i \in \{1, \dots, M\}$  and  $j \in \{2, \dots, N+1\}$ . Then, the log-likelihood is

$$\mathfrak{L}(\{V\}_{M,N}) = \sum_{i=1}^M \sum_{j=2}^{N+1} \log \left[ \rho_{i,j}(V_{i,j} | V_{i,j-1}) \right],$$

where  $\rho_{i,j}(V_{i,j} | V_{i,j-1}) = \rho_{i,j}(V_{i,j} | V_{i,j-1}; \xi_{1,i,j}, \xi_{2,i,j})$ , and where we assumed a non-informative prior.

# Data: CODE

We load our three tables: **Table\_Training\_Complete**, **Table\_Testing\_Complete**, and **Table\_Complete**.

```
1 function [Table_Training_Complete , Table_Testing_Complete , Table_Complete] = load_data(dataSet)
2
3 % 19/04/2020 15:45
4 % dataSet can be AWSTP, MTLOG or UTEP5.
5 load([' ../../ Python/Represas_Data_2/Wind_Data/' ,dataSet , '_0100_and_Real_24h_Training_Data.mat' ]);
6 load([' ../../ Python/Represas_Data_2/Wind_Data/' ,dataSet , '_0100_and_Real_24h_Testing_Data.mat' ]);
7 load([' ../../ Python/Represas_Data_2/Wind_Data/' ,dataSet , '_0100_and_Real_24h_Complete_Data.mat' ]);
8
9 %      Date           = Table_Training_Complete.Date;
10 %      Time           = Table_Training_Complete.Time;
11 %      Forecast       = Table_Training_Complete.Forecast;
12 %      Forecast_Dot   = Table_Training_Complete.Forecast_Dot;
13 %      real_ADME      = Table_Training_Complete.Real_ADME;
14 %      Error          = Table_Training_Complete.Error;
15 %      Error_Transitions = Table_Training_Complete.Error_Transitions;
16 %      Lamparti_Data   = Table_Training_Complete.Error_Lamp;
17 %      Lamparti-Tran   = Table_Training_Complete.Error_Lamp_Transitions;
18
19 end
```

From  $\theta_0$  to  $\theta_t$ :

To ensure that the analytical solutions is always in  $]0,1[$ , we choose the drift parameter to be

$$\theta(t) = \max \left( \theta_0, \frac{\theta_0 \alpha + |\dot{p}^\varepsilon(t)|}{\min(p^\varepsilon(t), 1 - p^\varepsilon(t))} \right), \quad \theta_0 > 0.$$

```
1 function [theta_t] = theta_t(theta_0, alpha, pe, pe_dot) % 13/03/2020 19:09
2
3     theta_t = max(theta_0, (theta_0*alpha + abs(pe_dot))/(min(pe,1-pe)));
4
5 end
```

Recall that we normalized w.r.t. time  $T$ . The normalized  $\hat{\theta}_t$  is  $\hat{\theta}_t = \theta_t T$ . Notice that in the definition of  $\theta_t$ , if we use  $\hat{\theta}_0 = \theta_0 T$  and the derivative of  $p_t$  w.r.t.  $d\hat{t}$  instead of w.r.t.  $dt$ , then we automatically have  $\hat{\theta}_t$ . See slide 5.



From  $p_t$  to  $p_t^\varepsilon$ :

To ensure that  $\theta(t)$  does not blow up, we truncate the forecast so it is always in  $[\varepsilon, 1 - \varepsilon]$  for some  $\varepsilon \in (0, 1/2)$ .

$$\text{Where } p^\varepsilon(t) = \begin{cases} \varepsilon & \text{if } p(t) < \varepsilon \\ p(t) & \text{if } \varepsilon \leq p(t) < 1 - \varepsilon \\ 1 - \varepsilon & \text{if } p(t) \geq 1 - \varepsilon \end{cases}.$$

```
1 function [pe] = pe(p, eps) % 18/03/2020 09:51
2     for i = 1:length(p)
3         if p(i) < eps
4             pe(i) = eps;
5         elseif p(i) >= eps && p(i) < 1-eps
6             pe(i) = p(i);
7         elseif p(i) >= 1-eps
8             pe(i) = 1-eps;
9         end
10    end
11 end
```

$\dot{p}_t^\varepsilon = \frac{dp^\varepsilon}{dt}$ . We use finite differences.

```
1 function [pe_dot] = pe_dot(pe, dt) % 18/03/2020 10:01
2
3     pe_dot = (pe(2:end) - pe(1:end-1)) / dt;
4
5 end
```

## Adapting the data:

We use this function to truncate the forecast, calculate its derivative, and to define the new errors ( $X_t - p_t^\varepsilon$ ) associated with the new truncated forecast.

We call to all this new set of data, the truncated

```
1 function [table_to_use] = load_data_eps(ep, dataSet, dataKind, dataNumber)
2
3 % 19/04/2020 15:45
4
5 if strcmp(dataKind, 'classic')
6     [Table_Training_Complete, Table_Testing_Complete, ~] = load_data(
7         dataSet);
8 elseif strcmp(dataKind, 'comparable')
9     [Table_Training_Complete, Table_Testing_Complete, ~] =
10         load_data_comparable(dataSet);
11 end
12
13 if strcmp(dataNumber, 'training')
14     table_to_use = Table_Training_Complete;
15 elseif strcmp(dataNumber, 'testing')
16     table_to_use = Table_Testing_Complete;
17 end
18
19 Time = table_to_use.Time;
20 Forecast = table_to_use.Forecast;
21 Real_ADME = table_to_use.Real_ADME;
22
23 dt = Time(1,2);
24 [M, ~] = size(Forecast);
25
26 for i = 1:M
27     Forecast(i,:) = pe(Forecast(i,:), ep);
28     Forecast_Dot(i,:) = pe_dot(Forecast(i,:), dt);
29     Error(i,:) = Real_ADME(i,:) - Forecast(i,:);
```

## Create a new batch (1/2):

This function is independent of if we are using the real data or the truncated one.

If we take a total of  $Z \in \mathbb{N}$  days, the batch corresponding to this days is

PATH 1							...	PATH Z
$t_n = 01 : 10$		$t_n = 01 : 20$		...	$t_n = 00 : 50$		...	...
$p(t_{n-1})$	$p(t_n)$	$p(t_{n-1})$	$p(t_n)$	...	$p(t_{n-1})$	$p(t_n)$		
$\dot{p}(t_{n-1})$	$\dot{p}(t_n)$	$\dot{p}(t_{n-1})$	$\dot{p}(t_n)$	...	$\dot{p}(t_{n-1})$	$\dot{p}(t_n)$		
$V(t_{n-1})$	$V(t_n)$	$V(t_{n-1})$	$V(t_n)$	...	$V(t_{n-1})$	$V(t_n)$		

with dimensions  $3 \times (2Z(N-1))$ . As an example: If we have 145 measurements ( $N+1$ ), then  $N = 144$  and  $N-1 = 143$ . We use 143 samples because we need to ignore the initial measurement (because we do not have data at time  $t_{-1}$ ) and the final one (because it does not have  $\dot{p}$ ). Then, each day has 143 samples. In this implementation, we are duplicating the data. In case of a lack of RAM, we can reduce the dimensions to  $3 \times (Z(N-1))$ .

## Create a new batch (2/2): CODE

```
1 function [Table_Training , new_bat] = new_batch_fixed( Table_Training , batch_size , N)
2     % 11/02/2020 09:35
3
4     Forecast      = Table_Training.Forecast;
5     Forecast_Dot  = Table_Training.Forecast_Dot;
6     Error         = Table_Training.Error;
7     new_bat       = [];
8
9     for i = 1:batch_size
10         for j = 2:N
11             forecast(2*j-3:2*j-2) = Forecast(i,j-1:j);
12             forecast_dot(2*j-3:2*j-2) = Forecast_Dot(i,j-1:j);
13             error(2*j-3:2*j-2) = Error(i,j-1:j);
14         end
15         new_bat = [new_bat , [forecast; forecast_dot; error]];
16     end
17 end
```

We call it **fixed** because, at some point, we were going to optimize the Likelihood taking random batches and increasing the size at each iteration. This idea was discarded since we have computational power enough to use all data at once.

## Complete batch:

This function is independent of if we are using the real data or the truncated one. After we created a batch with the elements  $(p(t), \dot{p}(t), V(t))$ , we want to add the parameter  $\theta(t)$  to use in the likelihood. Following the same idea than in slide 19, the complete batch is

PATH 1						...	PATH Z
$t_n = 01 : 10$		$t_n = 01 : 20$		...	$t_n = 00 : 50$		...
$p(t_{n-1})$	$p(t_n)$	$p(t_{n-1})$	$p(t_n)$	...	$p(t_{n-1})$	$p(t_n)$	
$\dot{p}(t_{n-1})$	$\dot{p}(t_n)$	$\dot{p}(t_{n-1})$	$\dot{p}(t_n)$	...	$\dot{p}(t_{n-1})$	$\dot{p}(t_n)$	
$V(t_{n-1})$	$V(t_n)$	$V(t_{n-1})$	$V(t_n)$	...	$V(t_{n-1})$	$V(t_n)$	
$\theta(t_{n-1})$	$\theta(t_n)$	$\theta(t_{n-1})$	$\theta(t_n)$	...	$\theta(t_{n-1})$	$\theta(t_n)$	

```
1 function [batch_theta] = batch_with_theta(batch, alpha, theta_0) % 09/02/2020 18:51
2     batch(4,1) = theta_t(theta_0, alpha, batch(1,1), batch(2,1));
3     batch(4,end) = theta_t(theta_0, alpha, batch(1,end), batch(2,end));
4     for i = 2:2:length(batch(1,:))-1
5         batch(4,i+1) = theta_t(theta_0, alpha, batch(1,i), batch(2,i));
6     end
7     batch_theta = batch;
8 end
```

## Initial guess for $\theta_0 \cdot \alpha$ : Quadratic variation

Recall we have  $M$  paths with  $N + 1$  measurements each. Depending on the SDE we are using, we have the two approximations:

$$X_t : \theta_0^* \alpha^* \approx \frac{1}{2M\Delta t} \sum_{j=1}^M \frac{\sum_{i=1}^N (\Delta X_{i,j})^2}{\sum_{i=1}^N (X_{i,j})(1 - X_{i,j})}, \quad V_t : \theta_0^* \alpha^* \approx \frac{1}{2M\Delta t} \sum_{j=1}^M \frac{\sum_{i=1}^N (\Delta V_{i,j})^2}{\sum_{i=1}^N (V_{i,j} + p_{i,j})(1 - V_{i,j} - p_{i,j})}$$

```
1 function [est] = initial_guess(real_prod, M, N, dt)
2     % 09/02/2020 09:30
3     est = 0;
4
5     for i = 1:M
6         numerator = 0; denominator = 0;
7         for j = 1:N
8             numerator = numerator + (real_prod(i, j+1) - real_prod(i, j))^2;
9             denominator = denominator + real_prod(i, j)*(1 - real_prod(i, j));
10        end
11        est = est + numerator/denominator;
12    end
13    est = est / (2*M*dt);
14
15 end
```

## Initial guess for $\theta_0$ (1/2): Mean least squares

Recall we have  $M$  paths with  $N + 1$  measurements each. If we assume  $\theta_t^\varepsilon = \theta_0$  for  $t \in [0, T]$ , then we have

$$\theta_0^* \approx \arg \min_{\theta_0} \left[ \sum_{j=1}^M \sum_{i=1}^N (V_{i+1,j} - V_{i,j}(1 - \theta_0 \Delta t))^2 \right]. \quad (1)$$

As problem (1) is convex, we can formulate the equivalent problem using derivatives

$$\theta_0^* \approx \frac{1}{\Delta t \cdot M} \sum_{j=1}^M \frac{\sum_{i=1}^N V_{i,j} (V_{i,j} - V_{i+1,j})}{\sum_{i=1}^N V_{i,j}^2}.$$

## Initial guess for $\theta_0$ (2/2): Mean least squares

As we need to assume that  $\theta_t^\varepsilon = \mathbf{c} \in \mathbb{R}^+$  in order to use correctly the estimator (1), we define  $\varepsilon, \gamma \in \mathbb{R}^+$ , and estimate  $\frac{\theta_0 \alpha}{\delta}$  using data which original forecast satisfies  $p_i \notin (\varepsilon, 1 - \varepsilon)$  and  $\theta_0$  using data which original forecast satisfies  $p_i \in [\gamma, 1 - \gamma]$ . The full explanation can be found in slides **20200308 - Initial Guess.pdf**.

```
1 function [val, accum] = mean_regression_eps(error, dt) % 16/03/2020 11:08
2     [M, N_ini] = size(error);
3     num = 0; den = 0;
4     accum = 0;
5     for i = 1:M
6         for j = 1:N_ini-1
7             if error(i, j+1) ~= -1
8                 num = num + error(i, j) * (error(i, j) - error(i, j+1));
9                 den = den + error(i, j)^2;
10                accum = accum + 1;
11            end
12        end
13    end
14    val = num/(den*dt);
15 end
```

To create this specific data sets, we use the functions **create\_samples\_eps.m** ( $\varepsilon$ -data) and **create\_samples\_minus\_eps.m** ( $\gamma$ -data).



## Initial guess for $\delta$ :

We have that, for almost all days, at time  $t = t_0$ ,  $X(t_0) \neq p(t_0)$  and then  $V(t_0) \neq 0$ . However, by forecast construction, there should exist a time  $t_{-\delta} < t_0$  such that  $V(t_{-\delta}) = 0$ .

We extrapolate linearly  $p(t)$  so we can evaluate  $p(t_{-\delta})$ , and assume that  $V(t_{-\delta}) = 0$ . Then, for each day  $j$ , we have an initial transition  $(V_{j,t_0} | V_{j,t_{-\delta}}; \boldsymbol{\theta}, \delta)$ . We assume again that it is Beta and apply the same moment matching as for the rest of transitions. With our initial guess for  $\boldsymbol{\theta}$ , we can construct our initial guess for  $\delta$  solving the problem

$$\delta \approx \arg \min_{\delta} \mathcal{L}_{\delta}(\boldsymbol{\theta}, \delta; V^{M,1}) = \arg \min_{\delta} \prod_{j=1}^M \rho_0(V_{j,t_0} | V_{j,t_{-\delta}}; \boldsymbol{\theta}, \delta).$$

See slides **20200319 - Time delta** for very detailed information.

We use  $\delta = 5.5\text{h}$ .

## Log-likelihood evaluation (1/2): CODE

```
1 function [value] = log_LH_evaluation(batch_complete, alpha, theta_0, dt) % 03/02/2020 19:42
2
3     for i = 1:length(batch_complete(1,:))/2
4
5         j = i*2; % This is the real index (parfor must go one-by-one).
6         % Recall that: j is t_n and j-1 is t_{n-1}.
7         p1 = batch_complete(1,j-1); p2 = batch_complete(1,j);
8         v1 = batch_complete(3,j-1); v2 = batch_complete(3,j);
9         th1 = batch_complete(4,j-1); th2 = batch_complete(4,j);
10        n = 100; % 50 discretizations for the ODEs.
11
12        m1 = moment_1(v1, theta_0, alpha, th1, th2, p1, p2, dt, n);
13        m2 = moment_2(v1, theta_0, th1, th2, p1, p2, alpha, m1, dt, n);
14        [xi1, xi2] = moments_matching(m1(end), m2(end));
15        if xi1 < 0 || xi2 < 0
16            val(i) = 0;
17        else
18            val(i) = log_dist(v2, xi1, xi2);
19        end
20
21    end
22
23    value = sum(val);
24
25 end
```

## Log-likelihood evaluation (2/2):

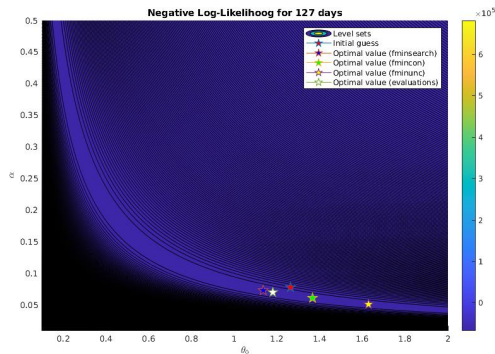


Figure 1: We use the code **plotLogLikelihood.m** to create this plots. We used about 18 thousand transitions, and we can also see the initial guess.

## Summing-up:

1. Initial guess: From quadratic variation and using all the data we get  $\hat{\theta}_0 \times \hat{\alpha}$ . Choosing some appropriate  $\gamma$ , from mean least squares we get  $\hat{\theta}_0$ , then we also get  $\hat{\alpha}$ . Also, from mean least squares we can estimate  $\varepsilon^*$ , using that the least squares estimator estimates  $\frac{\theta_0 \alpha}{\varepsilon}$  for some appropriate  $\varepsilon$ . Finally, using the pair  $(\hat{\theta}_0, \hat{\alpha})$ , we can obtain  $\hat{\eta}$  and  $\hat{\delta}$ .
2. Optimal parameters: Using the training data and the initial point  $(\hat{\theta}_0, \hat{\alpha})$ , we can find the optimal pair  $(\theta_0^*, \alpha^*)$  using the MLE. With this new pair, we can find  $\eta^*$  and  $\delta^*$ .

## Values for the parameters:

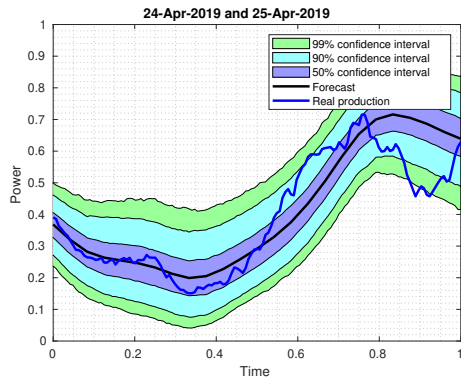
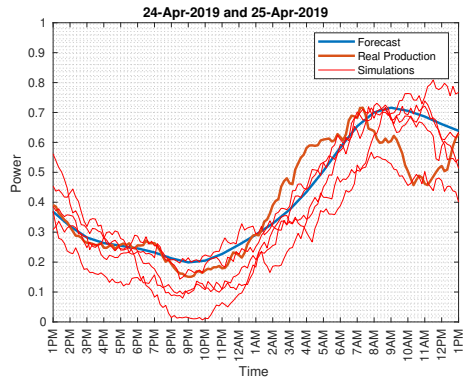
$\hat{\theta}_0$	$\hat{\alpha}$	$\hat{\eta}$	$\hat{\delta}$	$\gamma$	$\varepsilon$
1.63	0.06	*not apply*	220 min	0.3	0.018

$\eta^{ini}$  is not relevant here because we do not need to remove any data to find  $\delta^{ini}$ .

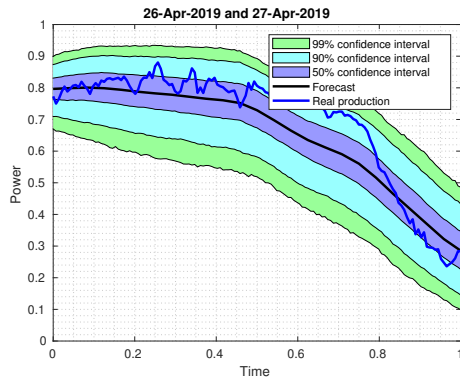
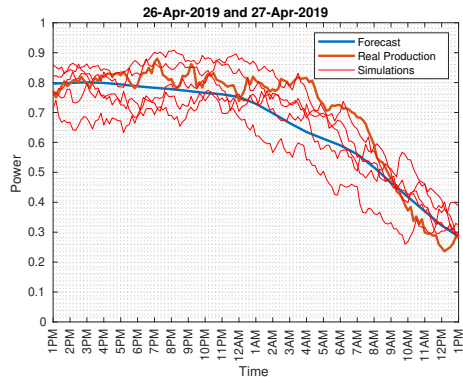
$\theta_0^*$	$\alpha^*$	$\eta^*$	$\delta^*$
3.91	0.02	0.1	220 min

Notice that  $\hat{\delta} = \delta^*$ . However, to obtain  $\delta^*$  we had to remove all initial errors larger than  $\eta^*$ , and in this way, we reduced the variance of the initial error. This has sense since  $\theta_0^* \alpha^* < \hat{\theta}_0 \hat{\alpha}$ .

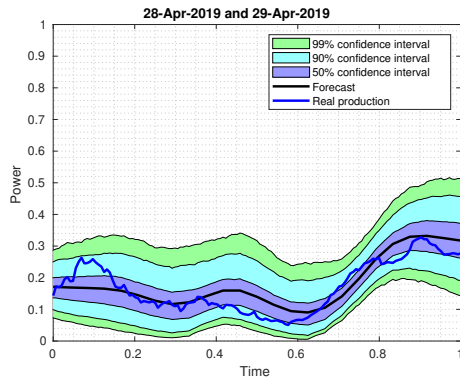
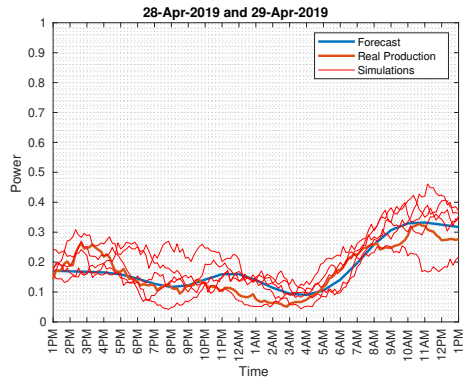
## Paths and bands for optimal values (1/4):



## Paths and bands for optimal values (2/4):

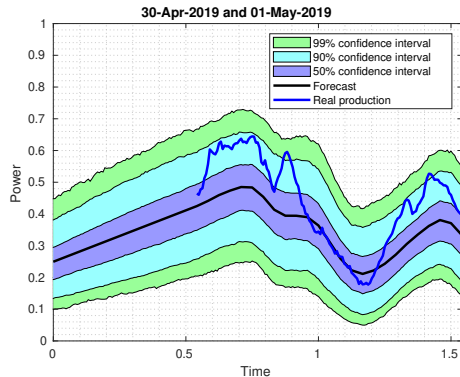
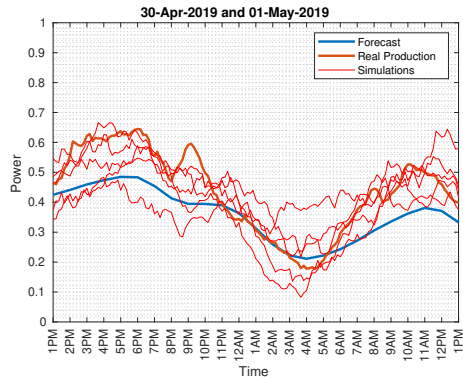


## Paths and bands for optimal values (3/4):





## Paths and bands for optimal values (4/4):



# Lamperti Transform

## Two candidates:

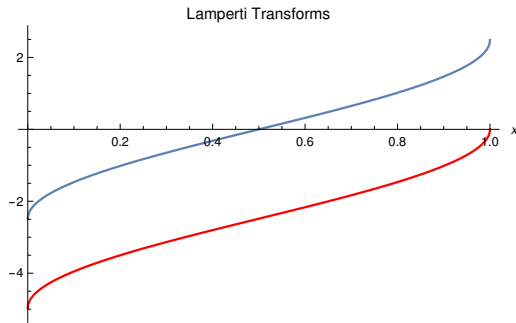
Recall  $X_t = V_t + p_t$ .

We have the candidates:

- ▶  $Z_t = \frac{1}{\sqrt{2\alpha\theta_0}} \arcsin(2(V_t + p_t) - 1)$ .
- ▶  $Z_t = -\sqrt{\frac{2}{\alpha\theta_0}} \arcsin(\sqrt{1 - V_t - p_t})$ .

They have the same partial derivatives. The main difference appears in the resulting SDE for  $Z_t$ .

For now, we will use the **second candidate**. The problem with the **first candidate** is that the mapping  $Z_t \rightarrow Z_t^2$  is not injective.



## Lamperti SDE: first candidate (this slide may be deleted so soon...)

The Lamperti SDE is:

$$dZ_t = \underbrace{\left[ \frac{(\alpha\theta_0 - \theta_t)\sin(\sqrt{2\alpha\theta_0}Z_t) - \theta_t(1 - 2p_t) + 2\dot{p}_t}{\sqrt{2\alpha\theta_0}\cos(\sqrt{2\alpha\theta_0}Z_t)} \right]}_{:=b(Z_t)} dt + dW_t.$$

Here we follow Bayesian Filtering and Smoothing, Chapter 9. We use the linearization-based approximation for the moments.

As  $X_t \in (0, 1)$ , we have that  $Z_t \in \left(-\frac{\pi}{2} \frac{1}{\sqrt{2\alpha\theta_0}}, \frac{\pi}{2} \frac{1}{\sqrt{2\alpha\theta_0}}\right)$ . Also, we have that (computed with Mathematica):

$$b'(z) = \frac{db}{dz}(z) = \frac{\alpha\theta_0 - \theta_t + (2\dot{p}_t + (2p_t - 1)\theta_t)\sin(z\sqrt{2\alpha\theta_0})}{\cos^2(z\sqrt{2\alpha\theta_0})}.$$

## Lamperti SDE: second candidate

The Lamperti SDE is:

$$dZ_t = \underbrace{\left[ \frac{\alpha\theta_0 \cos(Z_t\sqrt{2\alpha\theta_0}) - \theta_t \cos(Z_t\sqrt{2\alpha\theta_0}) + 2\theta_t p_t + 2\dot{p}_t - \theta_t}{\sqrt{\alpha\theta_0}\sqrt{1 - \cos(2Z_t\sqrt{2\alpha\theta_0})}} \right]}_{:=b(Z_t)} dt + dW_t.$$

Here we follow Bayesian Filtering and Smoothing, Chapter 9. We use the linearization-based approximation for the moments.

As  $X_t \in (0, 1)$ , we have that  $Z_t([0, 1]) = \left[ \frac{-\pi}{\sqrt{2\alpha\theta_0}}, 0 \right]$ . Also, we have that (computed with Mathematica):

$$b'(z) = \frac{db}{dz}(z) = -\frac{\alpha\theta_0 - \theta_t + \cos(\sqrt{2\alpha\theta_0}z)(2\dot{p}_t + \theta_t(2p_t - 1))}{\left(\sin(\sqrt{2\alpha\theta_0}z)\right)^2}.$$

## SDE first Lamperti moment (mean) (1/2):

Given some measurement  $z_{t_{n-1}}$ , we want to compute the first moment at time  $t_n$ . The linearly approximated first moment  $\mu_Z(s)$  for  $s \in [t_{n-1}, t_n]$  is the solution of the ODE

$$\begin{cases} d\mu_L(s) &= b(\mu_L(s))ds, \\ \mu_L(t_{n-1}) &= z_{t_{n-1}}. \end{cases}$$

We solve numerically the ODE using Forward-Euler:

$$\mu_L(s_n) = \mu_L(s_{n-1}) + b(\mu_L(s_{n-1}))\Delta s.$$

## SDE first Lamperti moment (mean) (2/2):

```
1 function m1 = moment_1_L(z, theta_0, alpha, p1, p2, dt, n)
2     % 29/03/2020 16:28
3     p_dot = (p2-p1) / dt;
4     m1(1) = z;
5     p_t = @(i) p1 + (p2-p1) * i/n;
6     ds = dt/n;
7
8     for i = 2:n
9         Theta_t = theta_t(theta_0, alpha, p_t(i), p_dot);
10        m1(i) = m1(i-1) + ds * ...
11            sde_Lamperti_drift_cand2(m1(i-1), alpha, theta_0, Theta_t, p_t(i-1), p_dot);
12    end
13
14 end
```

## SDE Lamperti variance (1/2):

Given some measurement  $z_{t_{n-1}}$ , we want to compute the variance at time  $t_n$ . The linearly approximated variance  $v_Z(s)$  for  $s \in [t_{n-1}, t_n]$  is the solution of the ODE

$$\begin{cases} d\sigma_L^2(s) &= (2\sigma_L^2 b'(\mu_L(s)) + 1) ds, \\ \sigma_L^2(t_{n-1}) &= 0. \end{cases}$$

We solve numerically the ODE using Forward-Euler:

$$\sigma_L^2(s_n) = \sigma_L^2(s_{n-1}) + \left(1 + 2\sigma_L^2(s_{n-1})b'(\mu_L(s_{n-1}))\right) \Delta s.$$



## SDE Lamperti variance (2/2):

```
1 function m2 = moment_2_L(m1, theta_0, alpha, p1, p2, dt, n)
2     % 29/03/2020 16:50
3
4     % This function returns the approximated variance.
5
6     p_dot = (p2-p1) / dt;
7     m2(1) = 0;
8     p_t = @(i) p1 + (p2-p1) * i/n;
9     ds = dt/n;
10
11     for i = 2:n
12         Theta_t = theta_t(theta_0, alpha, p_t(i), p_dot);
13         b_prime = sde_Lamperti_drift_prime_cand2(m1(i-1), alpha, theta_0, Theta_t, p_t(i-1), p_dot);
14         m2(i) = m2(i-1) + ds * ...
15             (1+2*m2(i-1)*b_prime);
16     end
17
18 end
```

## Density next measurement:

We want the next measurement  $Z_{t_n}|Z_{t_{n-1}}$  to have a Gaussian with mean  $\mu_Z$  and variance  $\sigma_Z^2$  ( $f_Z(z) = \mathcal{N}(\mu_Z, \sigma_Z^2)$ ).

Then, we want the SDE and our new PDF  $f_Z(z)$  to have the same moments at each  $t \in \{\text{some appropriate domain}\}$ , i.e.,  $\mu_Z(t) = \mu_L(t)$  and  $\sigma_Z^2(t) = \sigma_L^2(t)$ .

The density and log-density (natural logarithm) are:

$$f_Z(z) = \frac{e^{-\frac{1}{2}\left(\frac{z-\mu_Z}{\sigma_Z}\right)^2}}{\sigma_Z\sqrt{2\pi}} \iff \log(f_Z(z)) = -\frac{1}{2}\left(\frac{z-\mu_Z}{\sigma_Z}\right)^2 - \log(\sigma_Z\sqrt{2\pi}).$$

```
1 function [val] = log_dist_L(z,mu,sig) % 29/03/2020 17:49
2
3     val = -(1/2)*((z-mu)/sig)^2 - log(sig*sqrt(2*pi));
4
5 end
```

# Lamperti Log-likelihood evaluation: CODE

```
1 function [value] = log_LH-evaluation-L(batch_complete, theta0, alpha, dt) % 03/02/2020 19:42
2
3     for i = 1:length(batch_complete(1,:))/2
4
5         j = i*2; % This is the real index (parfor must go one-by-one).
6         % Recall that: j is t_n and j-1 is t_{n-1}.
7         p1 = batch_complete(1,j-1); p2 = batch_complete(1,j);
8         z1 = batch_complete(5,j-1); z2 = batch_complete(5,j);
9         n = 50; % 10 discretizations for the ODEs.
10
11         mu      = moment_1_L(z1, theta0, alpha, p1, p2, dt, n);
12         sig2     = moment_2_L(mu, theta0, alpha, p1, p2, dt, n);
13         sig      = sqrt(sig2(end));
14         val(i) = log_dist_L(z2, mu(end), sig);
15
16     end
17
18     value = sum(val);
19
20 end
```

# Model 1

## Model 1 SDE:

Given the SDE

$$dX_t = \underbrace{-\theta_0 (X_t - p_t)}_{:=\mu_t} dt + \underbrace{\sqrt{2\theta_0 \alpha X_t (1 - X_t)}}_{:=\sigma_t} dW_t,$$

we want to approximate its moments. Trivially, we have an exact ODE for the first moment for  $t \in [t_0, t_1]$ :

$$\begin{cases} d\mathbb{E}[X_t] = -\theta_0 (\mathbb{E}[X_t] - p_t) dt \\ \mathbb{E}[X_{t_0}] = x_{t_0}. \end{cases}$$

Now, we can define the process  $f_t = f(X_t, t) = X_t^2$  and apply Itô's lemma. We reach that

$$df_t = \left( \frac{\partial f}{\partial t} + \mu_t \frac{\partial f}{\partial x} + \frac{\sigma_t^2}{2} \frac{\partial^2 f}{\partial x^2} \right) dt + \sigma_t \frac{\partial f}{\partial x} dW_t,$$

where  $\frac{\partial f}{\partial x} = 2x$ ,  $\frac{\partial^2 f}{\partial x^2} = 2$ , and  $\frac{\partial f}{\partial t} = 0$ .

## Model 1 SDE:

Doing SOME substitutions, we reach

$$df_t = 2\theta_0 \left( X_t(\alpha + p_t) - X_t^2(1 + \alpha) \right) dt + \sigma_t \frac{\partial f}{\partial x} dW_t.$$

Now, if we apply the expected value operator  $\mathbb{E}[\cdot]$ , we get

$$d\mathbb{E} \left[ X_t^2 \right] = 2\theta_0 \left( \mathbb{E} [X_t] (\alpha + p_t) - \mathbb{E} \left[ X_t^2 \right] (1 + \alpha) \right) dt,$$

which is the second moment ODE with IV  $\mathbb{E} \left[ X_t^2 \right] = x_{t_0}^2$  for  $t \in [t_0, t_1]$ . Recall that  $\mathbb{V} [X_t] = \mathbb{E} \left[ X_t^2 \right] - \mathbb{E} [X_t]^2$ . As we will assume Gaussian proxies, we will use the Likelihood described in Slide [42](#).

## Model 1 SDE:

The forward Euler schemes are:

1.  $m_1(s_n) = m_1(s_{n-1}) - \theta_0(m_1(s_{n-1}) - p_{s_{n-1}})\Delta s.$
2.  $m_2(s_n) = m_2(s_{n-1}) + 2\theta_0(m_1(s_{n-1})(\alpha + p_{s_{n-1}}) - m_2(s_{n-1})(1 + \alpha))\Delta s.$

```
1 function m1 = moment_1.M1(x,theta_0,p1,p2,dt,n) % 19/04/2020 21:56
2     p = @(i) p1 + (p2-p1) * i/n; % We interpolate p(t).
3     m1(1) = x;
4     ds = dt/n;
5     for i = 2:n
6         m1(i) = m1(i-1) - theta_0*(m1(i-1)-p(i-1))*ds;
7     end
8 end
```

```
1 function m2 = moment_2.M1(x,m1,theta_0,alpha,p1,p2,dt,n) % 19/04/2020 22:00
2     p = @(i) p1 + (p2-p1) * i/n; % We interpolate p(t).
3     m2(1) = x^2;
4     ds = dt/n;
5     for i = 2:n
6         m2(i) = m2(i-1) + 2*theta_0*(m1(i-1)*(alpha+p(i-1)) ...
7             - m2(i-1)*(1+alpha))*ds;
8     end
9
10 end
```

## Model 1 SDE: Shoji-Ozaki method

An extension of the method to Ozaki the more general case where the drift is allowed to depend on the time variable  $t$ , and also the diffusion coefficient can be varied is the method Shoji and Ozaki. Consider the stochastic differential equation:

$$dX_t = f(t, X_t, \boldsymbol{\theta})dt + q(X_t, \boldsymbol{\theta})dW_t, \quad t \geq 0, X_0 = x_0,$$

the transition density for the Shoji-Ozaki method is Gaussian, we have that:

$X_{t+\Delta t} | X_t = x \sim \mathcal{N} \left( A_{(t,x)}x, B_{(t,x)}^2 \right)$ , where:

$$A_{(t,x)} = 1 + \frac{f(t,x)}{xL_t} \left( e^{L_t\Delta t} - 1 \right) + \frac{M_t}{xL_t^2} \left( e^{L_t\Delta t} - 1 - L_t\Delta t \right),$$

$$B_{(t,x)} = g(x) \sqrt{\frac{e^{2L_t\Delta t} - 1}{2L_t}},$$

with:

$$L_t = \partial_x f(t, x) \quad \text{and} \quad M_t = \frac{g^2(x)}{2} \partial_{xx} f(t, x) + \partial_t f(t, x).$$

<https://cran.r-project.org/web/packages/Sim.DiffProc/vignettes/fitsde.html>



## Model 1 SDE: Beta as proxy

Following the same idea as in Slides (10) and (11), but using  $[a, b] = [0, 1]$ , we get the moments matching:

- ▶  $\xi_1 = \frac{-\mu\sigma^2 + \mu^2 - \mu^3}{\sigma^2}$  all evaluated at time  $t_n$  (verified in **Mathematica 11.0**<sup>2</sup>).
- ▶  $\xi_2 = \frac{(\mu-1)(\sigma^2 - \mu + \mu^2)}{\sigma^2}$  all evaluated at time  $t_n$  (verified in **Mathematica 11.0**).

---

<sup>2</sup>File: [Model1\\_Beta.nb](#).