Homework 4

Xingchen Yu, Meng Rui and Alex Terenin

February 24, 2017

Problem 2.21

(21) Perform analyses of the USA/UK exchange rate index series along the lines of those in Section 2.6, one for each value of the discount factor $\delta = 0.6, 0.65, \ldots, 0.95, 1$. Relative to the DLM with $\delta = 1$, plot the MSE, MAD and LLR measures as functions of δ. Comment on these plots. Sensitivity analyses explore how inferences change with respect to model assumptions. At $t = 115$, explore how sensitive this model is to values of δ in terms of inferences about the final level μ_{115}, the variance V and the next observation Y_{116}.

Table 1.USA/UK exchange rate index (×100)

Year	Month (Jan – Jun) & (Jul – Dec)					
75	1.35	1.00	−1.96	−2.17	−1.78	−4.21
	−3.30	−1.43	−1.35	−0.34	−1.38	0.30
76	−0.10	−4.13	−5.12	−2.13	−1.17	−1.24
	1.01	−3.02	−5.40	−0.12	2.47	2.06
77	−0.18	0.29	0.23	0.00	0.06	0.17
	0.98	0.17	1.59	2.62	1.96	4.28
78	0.26	−1.66	−3.03	−1.80	1.04	3.06
	2.50	0.87	2.42	−2.37	1.22	1.05
79	−0.05	1.68	1.70	−0.73	2.59	6.77
	−0.98	−1.71	−2.53	−0.61	3.14	2.96
80	1.01	−3.69	0.45	3.89	1.38	1.57
	−0.08	1.30	0.62	−0.87	−2.11	2.48
81	−4.73	−2.70	−2.45	−4.17	−5.76	−5.09
	−2.92	−0.22	1.42	3.26	0.05	−0.95
82	−2.14	−2.19	−1.96	2.18	−2.97	−1.89
	0.12	−0.76	−0.94	−3.90	−0.86	−2.88
83	−2.58	−2.78	3.30	2.06	−1.54	−1.30
	−1.78	−0.13	−0.20	−1.35	−2.82	−1.97
84	2.25	1.17	−2.29	−2.49	−0.87	−4.15
	−0.53					

Time Series Plot of the Exchange Rate Index

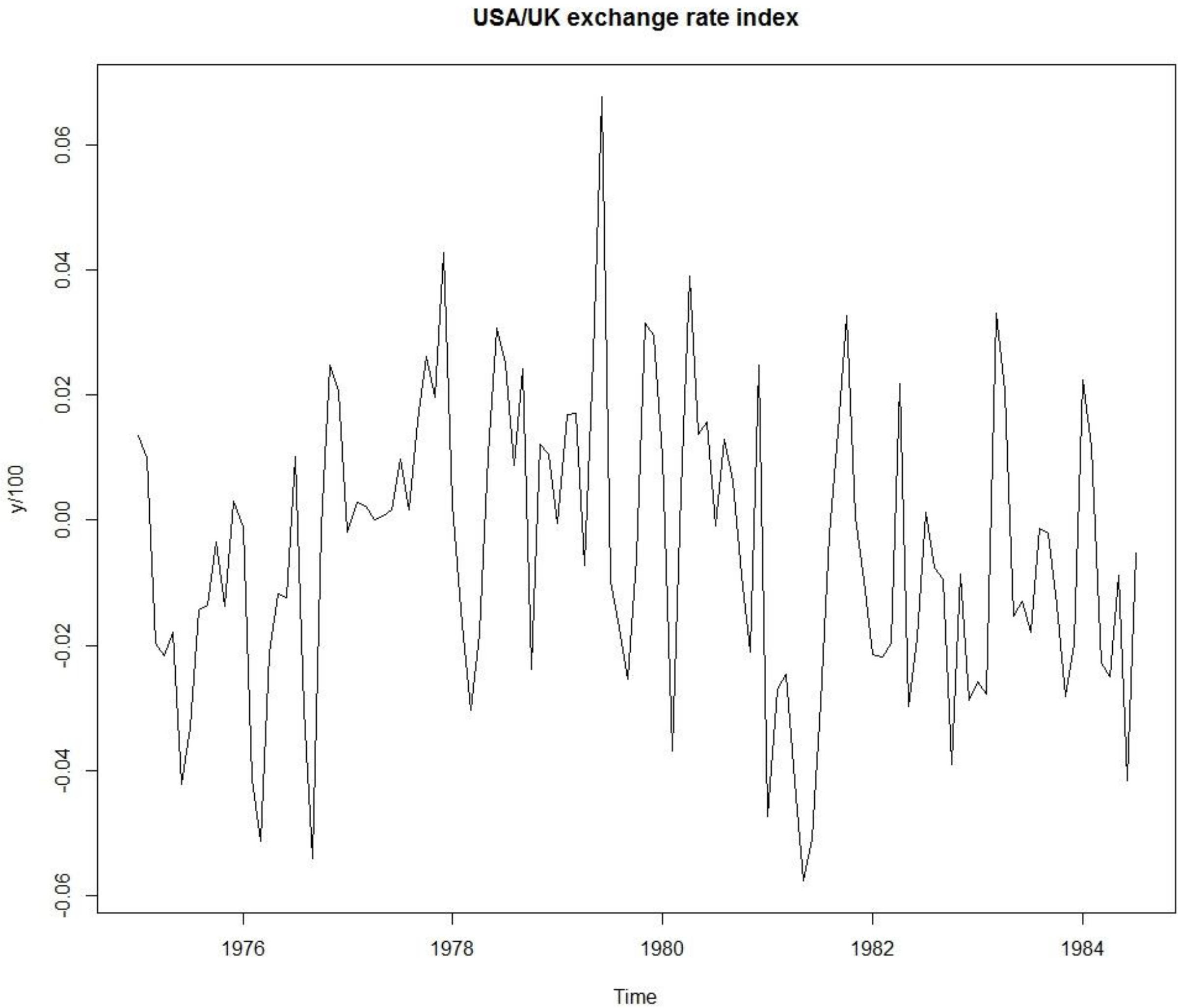


Figure 1. USA/UK exchange rate index

Model Specification (Unknown Observational Variances)

First Order Polynomial DLM (1, 1, V, V$\boldsymbol{W_t}$)

$$Y_t = \mu_t + \nu_t,\ \ \nu_t \sim N(0, V)$$

$$\mu_t = \mu_{t-1} + \omega_t,\ \ \omega_t \sim N(0, VW_t)$$

$$where\ \ W_t = \frac{(1-\delta)}{\delta} C_{t-1}\ and\ \delta\ is\ discount\ factor, \delta \in (0,1]$$

Prior Distribution

$$(V \mid D_0) \sim IG\left(\frac{n_0}{2}, \frac{d_0}{2}\right)$$

$$(\mu_0 \mid D_0, V) \sim N(m_0, VC_0)$$

Recursive Update Equations Conditioned on V

$$Posterior\ at\ time\ t-1:\quad (\mu_{t-1} \mid D_{t-1}, V) \sim N(m_{t-1}, VC_{t-1})$$

$$(\nu_{t-1} \mid D_{t-1}) \sim IG\left(\tfrac{n_{t-1}}{2}, \tfrac{d_{t-1}}{2}\right)$$

$$Prior\ at\ time\ t:\quad (\mu_t \mid D_{t-1}, V) \sim N(m_{t-1}, VR_t)$$

$$(\nu_t \mid D_t) \sim IG\left(\tfrac{n_t}{2}, \tfrac{d_t}{2}\right)$$

$$One\ step\ ahead\ predictive:\quad (Y_t \mid D_{t-1}, V) \sim N(f_t, VQ_t)$$

With Recursive Parameter Update

$$R_t = C_{t-1} + W_t$$

$$f_t = m_{t-1}$$

$$Q_t = R_t + 1$$

$$e_t = Y_t - f_t$$

$$A_t = R_t / Q_t$$

$$C_t = R_t - A_t^2 Q_t$$

$$m_t = m_{t-1} + A_t e_t$$

$$n_t = n_{t-1} + 1$$

$$d_t = d_{t-1} + \epsilon_t^2 / Q_t$$

Marginalized Over Unknown Variance V

$$Posterior\ at\ time\ t-1: \quad (\mu_{t-1} \mid D_{t-1}) \sim T_{n_{t-1}}(m_{t-1}, C_{t-1})$$

$$Prior\ at\ time\ t: \quad (\mu_t \mid D_{t-1}) \sim T_{n_{t-1}}(m_{t-1}, R_t)$$

$$One\ step\ ahead\ predictive: \quad (Y_t \mid D_{t-1}) \sim T_{n_{t-1}}(f_t, Q_t)$$

With Recursive Parameter Update

$$R_t = C_{t-1} + W_t$$

$$Q_t = R_t + S_{t-1}$$

$$A_t = R_t / Q_t$$

$$f_t = m_{t-1}$$

$$e_t = Y_t - f_t$$

$$m_t = m_{t-1} + A_t e_t$$

$$n_t = n_{t-1} + 1$$

$$d_t = d_{t-1} + S_{t-1}\epsilon_t^2 / Q_t$$

$$S_t = d_t / n_t$$

$$C_t = A_t S_t$$

Initial Prior distribution (vague, uninformative) is defined by

$$m_0 = 0, C_0 = 1, n_0 = 1\ and\ d_0 = 0.01$$

Such that μ_0 lies between -0.1 and 0.1 with probability 0.5 and -0.63 and 0.63 with probability 0.9.

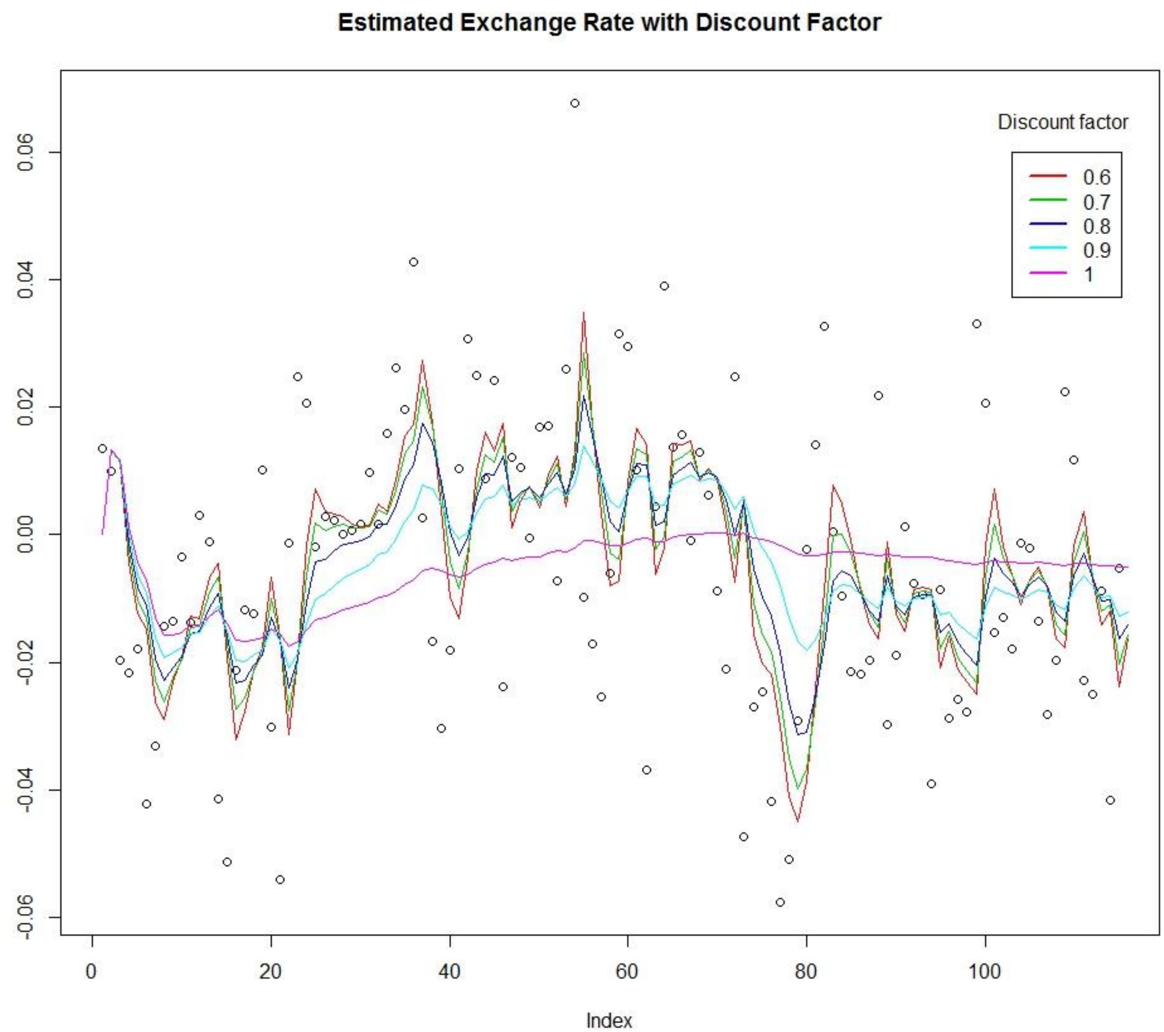


Figure 2. Estimated Exchange Rate with Discount Factor

Model Evaluation

$$MAD = \frac{1}{115}\sum_{t=1}^{t=115} |e_t|, MSE = \frac{1}{115}\sum_{t=1}^{t=115} e_t^2, LLR(\delta) = L(\delta) - L(1)$$

$$L(\delta) = \log\big(p(y_{1:115} \,|D_0, \delta)\big) = \sum_{t=1}^{t=115} \log(p(\, y_t \mid D_{t-1}, \delta)$$

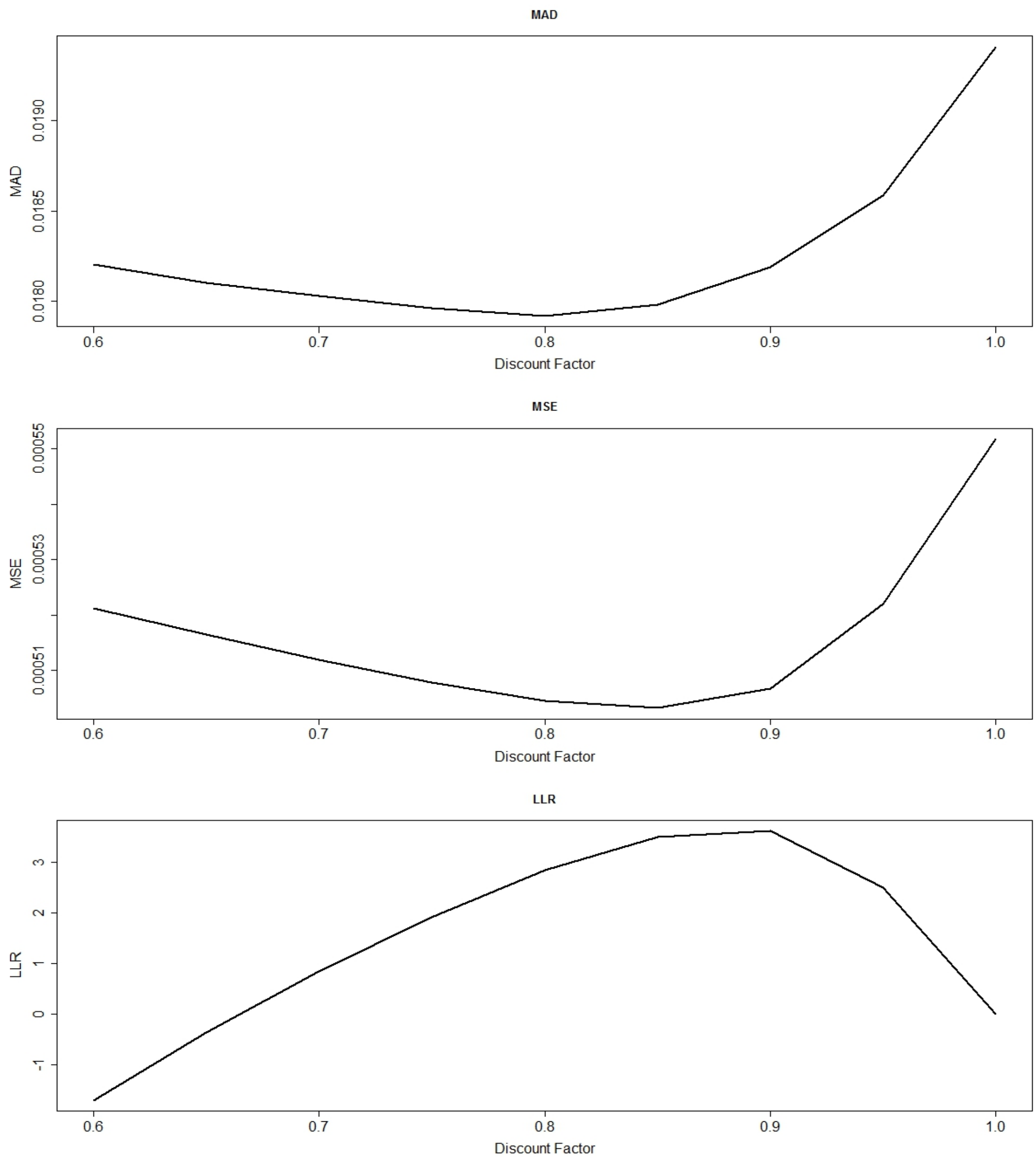


Figure 3. Model Evaluation via MAD, MSE and LLR

Sensitivity Analysis

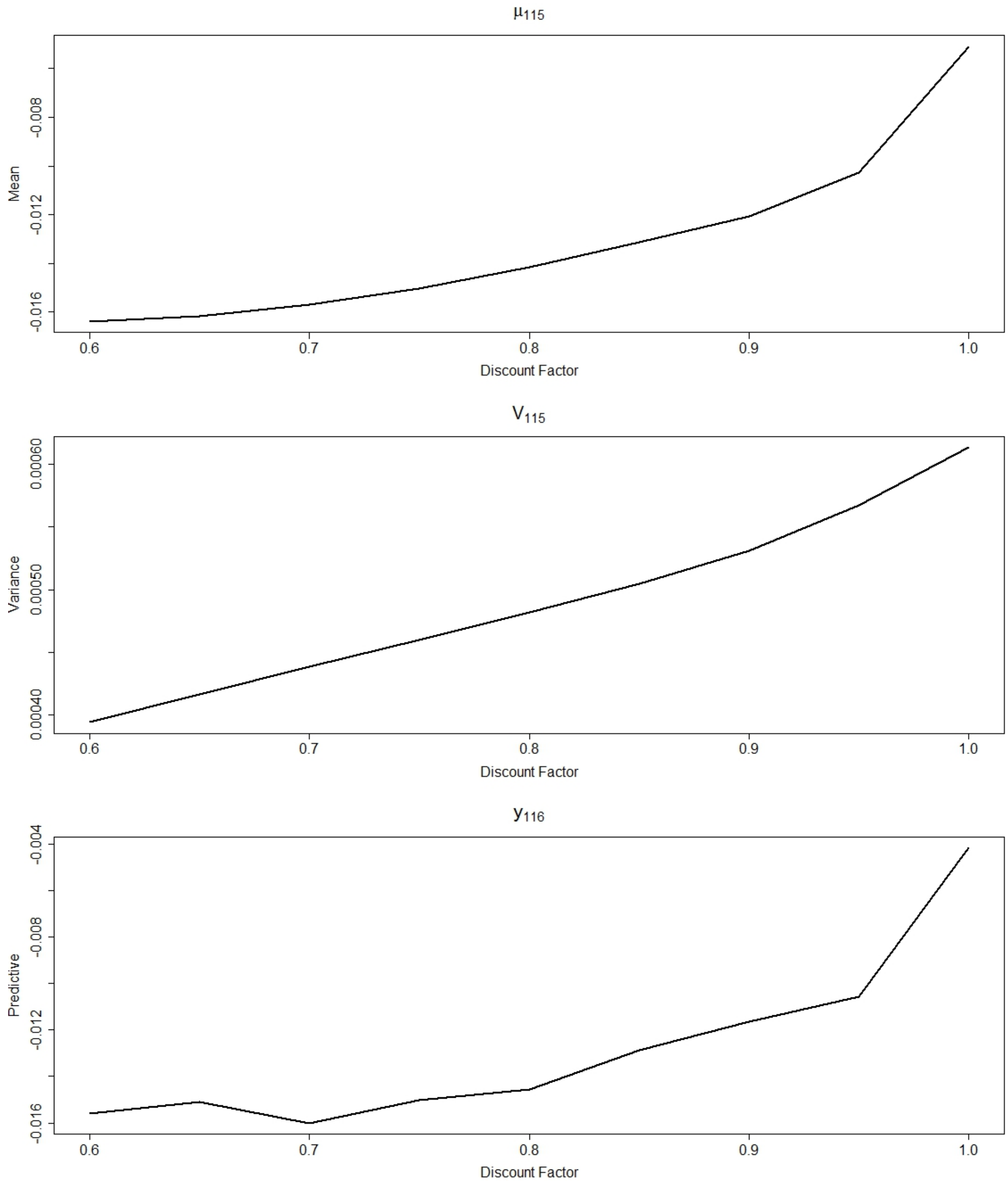


Figure 4. Sensitivity Analysis of Mean and Variance at y_{115} and one step forecast at y_{116}

Recursive Update Function with Discount Factor

```
##############################################

update.discount<-function(Y,delta,m.0,C.0,n.0,S.0){

 N <- length(Y)+1

 W<-n<-d<-m<-C<-R<-Q<-S<-f<-A<-e<-rep(0,length=N)

 Y<-c(0,Y)

 ##############################################

 C[1]<-C.0

 m[1]<-m.0

 S[1]<-S.0

 n[1]<-n.0

 ##############################################

 for ( t in 2:N ) {

  n[t] <- n[t-1] + 1

  W[t] <- C[t-1] * (1-delta) / delta

  R[t] <- C[t-1] + W[t]

  f[t] <- m[t-1]

  Q[t] <- R[t] + S[t-1]

  A[t] <- R[t] / Q[t]

  e[t] <- Y[t] - f[t]

  S[t] <- S[t-1]+(S[t-1]/n[t])*(e[t]^2/Q[t]-1)

  m[t] <- m[t-1]+A[t]*e[t]

  C[t] <- A[t]*S[t]

 }

 return (list(m=m,C=C,R=R,f=f,Q=Q,S=S))

}

##############################################
```

Question 4.1

February 23, 2017

1 Question:

In Chapter 2, we consider random walk-plus-noise model for the Nile river data. There, we used the maximum likelihood estimates for the state and observation variances. Consider now Bayesian inference on the states and unknown parameters of the model. Express conjugate priors for V and W and evaluate the posterior distribution of $(\theta_{0:T}, v, w|y_{1:T})$. Then, estimate the model using discount factors as in Sections 4.3.2 and 4.3.3, and compare the results.

2 Solution

Model:

$$Y_t = F_t\theta_t + v_t \quad v_t \sim \mathcal{N}(0, v_t) \tag{1}$$

$$\theta_t = G_t\theta_t + \omega_t \quad \omega_t \sim \mathcal{N}(0, w_t) \tag{2}$$

Let $v_t = v = \phi_y^{-1}$ and $w_t = w = \phi_\theta^{-1}$ and both of them have independent inverse Gamma distributions, which implies (ϕ_y, ϕ_θ) has the prior that is the product of tow Gamma density.
Donate $E(\phi_y) = a_y$ and $E(\phi_\theta) = a_\theta$ with variance $V(\phi_y) = b_y$ and $V(\phi_\theta) = b_\theta$. Thus the Gamma priors can be parameterized as

$$\phi_y \sim \mathcal{G}(\alpha_y, \beta_y) \quad \phi_\theta \sim \mathcal{G}(\alpha_\theta, \beta_\theta)$$

with $\alpha_y = \frac{a_y^2}{b_y}$, $\beta_y = \frac{a_y}{b_y}$, $\alpha_\theta = \frac{a_\theta^2}{b_\theta}, \beta_\theta = \frac{a_\theta}{b_\theta}$.
It also means that

$$v \sim \mathcal{IG}(\alpha_y, \beta_y) \quad w \sim \mathcal{IG}(\alpha_\theta, \beta_\theta)\,.$$

Given the observations $y_{1:T}$, the joint posterior of the states $\theta_{0:T}$ and the unknown parameters $\psi = (\theta_y, \theta_\theta)$ is proportional to the joint density.

$$\begin{aligned}
\pi(\theta_{0:T}, \psi|y_{1:T}) &\propto \pi(\theta_{0:T}, \psi, y_{1:T}) && (3)\\
&= \pi(y_{1:T}|\theta_{0:T}, \psi)\pi(\theta_{0:T}|\psi)\pi(\psi) && (4)\\
&= \Pi_{t=1}^{T}\pi(y_t|\theta_t, \phi_y)\Pi_{t=1}^{T}\pi(\theta_t|\theta_{t-1}, \phi_\theta)\pi(\theta_0)\pi(\phi_y)\pi(\phi_\theta) && (5)
\end{aligned}$$

Also we can rewrite as

$$
\begin{aligned}
\pi(\theta_{0:T}, v, w|y_{1:T}) &\propto \pi(\theta_{0:T}, v, w, y_{1:T}) && (6)\\
&= \pi(y_{1:T}|\theta_{0:T}, v, w)\pi(\theta_{0:T}|v, w)\pi(v, w) && (7)\\
&= \Pi_{t=1}^{T}\pi(y_t|\theta_t, v)\Pi_{t=1}^{T}\pi(\theta_t|\theta_{t-1}, w)\pi(\theta_0)\pi(v)\pi(w) && (8)\\
&= \Pi_{t=1}^{T}\mathcal{N}(y_t|F_t\theta_t, v)\Pi_{t=1}^{T}\mathcal{N}(\theta_t|G_t\theta_{t-1}, w)\pi(\theta_0)\mathcal{IG}(\alpha_y, \beta_y)\mathcal{IG}(\alpha_\theta, \beta_\theta) && (9)
\end{aligned}
$$

For the computation convenience, we compute $\pi(\phi_y|y_{1:T}, \theta_{0:T}, \phi_\theta)$ and $\pi(\phi_\theta|y_{1:T}, \theta_{0:T}, \phi_y)$.

$$
\begin{aligned}
\pi(\phi_y|y_{1:T}, \theta_{0:T}, \phi_\theta) &\propto \pi(\theta_{0:T}, \psi|y_{1:T}) && (10)\\
&\propto \Pi_{t=1}^{T}\pi(y_t|\theta_t, \phi_y)\pi(\phi_y) && (11)\\
&\propto \phi_y^{\frac{T}{2}+\alpha_y-1}e^{-\phi_y[\frac{1}{2}\sum_{i=1}^{T}(y_t-F_t\theta_t)^2+\beta_y]} && (12)
\end{aligned}
$$

which implies $\pi(\phi_y|y_{1:T}, \theta_{0:T}, \phi_\theta) \sim \mathcal{G}(\alpha_y+\frac{T}{2}, \beta_y+\frac{1}{2}SS_y)$ with $SS_y = \sum_{i=1}^{T}(y_t - F_t\theta_t)^2$ and it is independent of ϕ_θ. Therefore

$$
\pi(\phi_y|y_{1:T}, \theta_{0:T}) \sim \mathcal{G}(\alpha_y + \frac{T}{2}, \beta_y + \frac{1}{2}SS_y)\,.
$$

Also,

$$
\begin{aligned}
\pi(\phi_\theta|y_{1:T}, \theta_{0:T}, \phi_y) &\propto \pi(\theta_{0:T}, \psi|y_{1:T}) && (13)\\
&\propto \Pi_{t=1}^{T}\pi(\theta_t|\theta_{t-1}, \phi_\theta)\pi(\phi_\theta) && (14)
\end{aligned}
$$

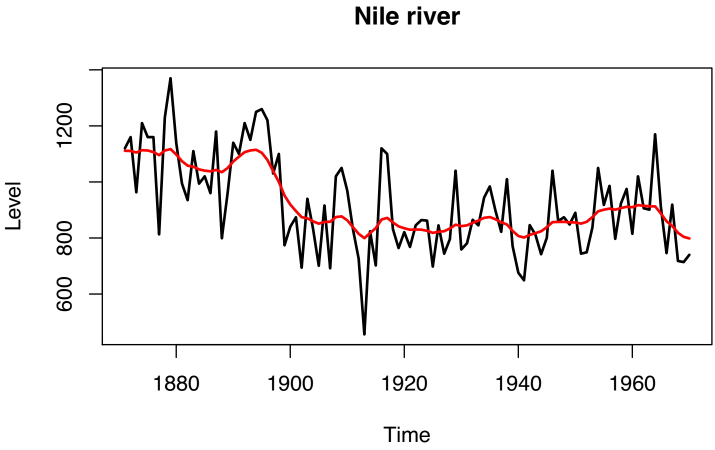
Then following the same approach, we have $\pi(\phi_\theta|y_{1:T}, \theta_{0:T}) \sim \mathcal{G}(\alpha_\theta + \frac{T}{2}, \beta_\theta + \frac{1}{2}SS_\theta)$ with $SS_\theta = \sum_{t=1}^{T}(\theta_t - G_t\theta_{t-1})^2$. After re-parameterization, we have

$$
\pi(v|y_{1:T}, \theta_{0:T}) \sim \mathcal{IG}(\alpha_y + \frac{T}{2}, \beta_y + \frac{1}{2}SS_y) \quad (15)
$$

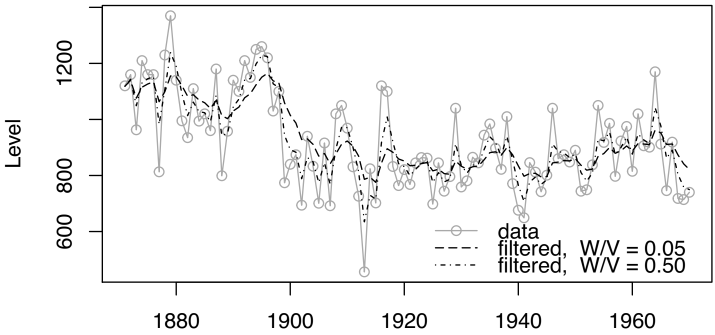
$$
\pi(w|y_{1:T}, \theta_{0:T}) \sim \mathcal{IG}(\alpha_\theta + \frac{T}{2}, \beta_\theta + \frac{1}{2}SS_\theta) \quad (16)
$$

which show the priors of v and w are conjugate.

Nile River Data

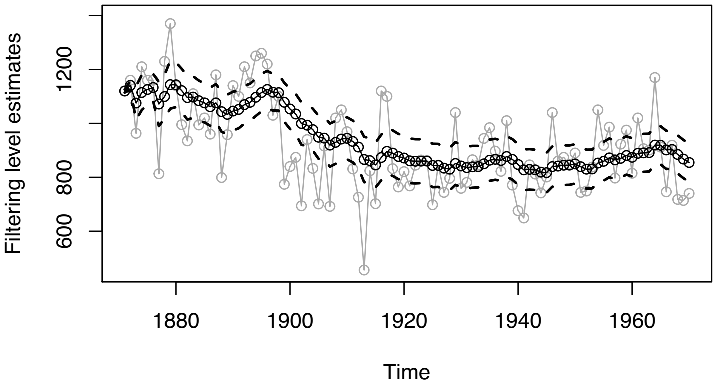


Fixed W,V

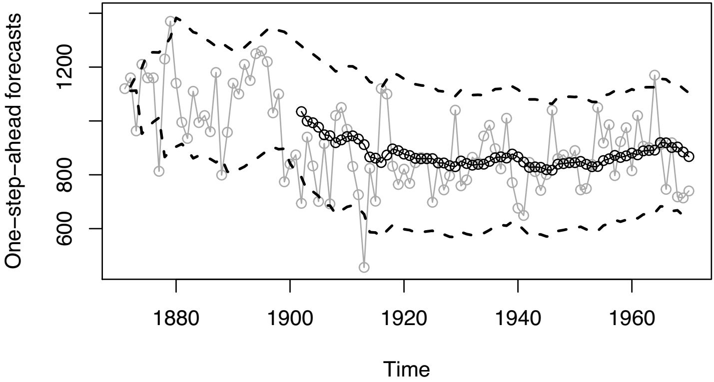


Level
1200
800
600
1880
1900
1920
1940
1960
data
filtered, W/V = 0.05
filtered, W/V = 0.50

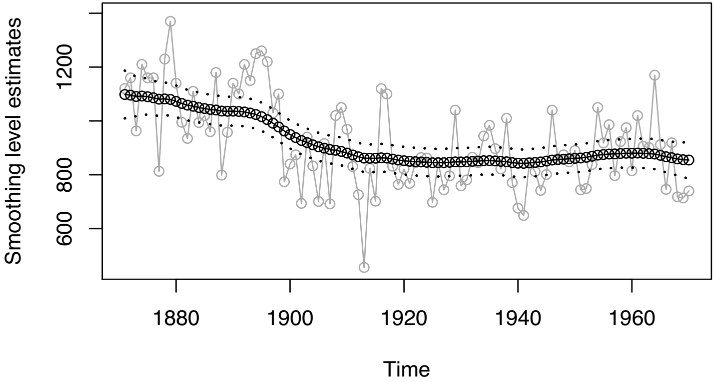
Filtering: $\delta = 0.9$



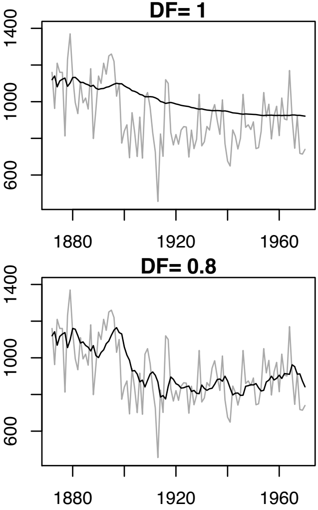
1-step ahead forecast: $\delta = 0.9$



Smoothing: $\delta = 0.9$



1-step ahead forecast comparison: $\delta \in \{1, 0.9, 0.8, 0.3\}$



1-step ahead forecast comparison: $\delta \in \{1, 0.9, 0.8, 0.3\}$

DF	MAPE	MAD	MSE	SD
1.0	0.1819	150.8595	41988.91	27796.047
0.9	0.1439	125.2313	34016.08	18580.757
0.8	0.1398	123.3994	33197.85	15991.467
0.3	0.1461	131.0214	35455.36	6728.878

1-step ahead forecast: time-varying V_t