VARX

Sri Santhosh Hari

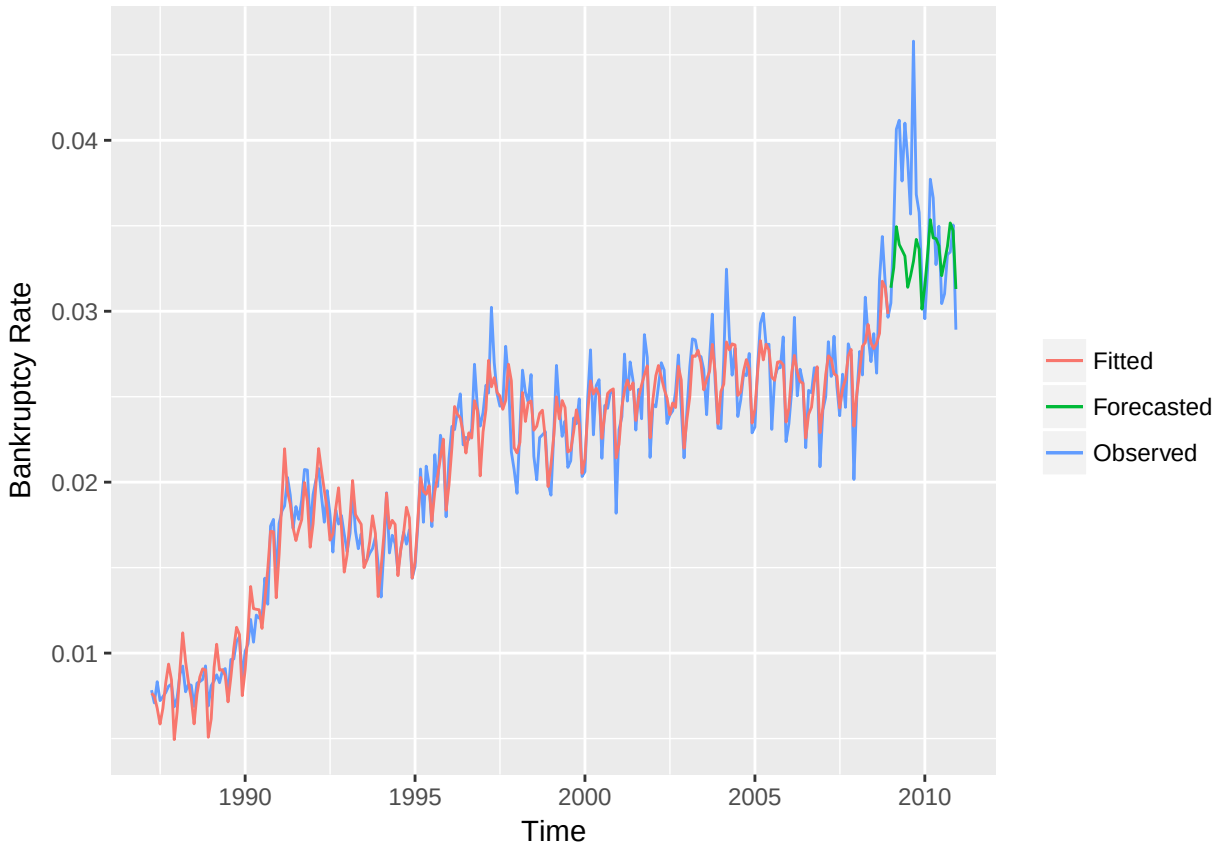
```
train_data <- read_csv("Data/train.csv")
train_data <- na.omit(train_data)
test <- read_csv("Data/test.csv")
test <- na.omit(test)
train_data$Mon <- seq.Date(as.Date("1987/1/1"), as.Date("2010/12/1"), by = "month")
test$Mon <- seq.Date(as.Date("2011/1/1"), as.Date("2012/12/1"), by = "month")
train <- train_data[1:264,]
valid <- train_data[265:288,]
```

```
VARselect(train[c("Bankruptcy_Rate","Unemployment_Rate","House_Price_Index")],
          lag.max = 12, type="both", season=12, exogen=train["Population"])
```

```
## $selection
## AIC(n)  HQ(n)  SC(n) FPE(n)
##      3      3      3      3
##
## $criteria
##                   1             2             3             4
## AIC(n) -1.846788e+01 -1.890259e+01 -1.912562e+01 -1.909761e+01
## HQ(n)  -1.818046e+01 -1.856446e+01 -1.873677e+01 -1.865804e+01
## SC(n)  -1.775359e+01 -1.806225e+01 -1.815923e+01 -1.800517e+01
## FPE(n)  9.544891e-09  6.182210e-09  4.948935e-09  5.092983e-09
##                   5             6             7             8
## AIC(n) -1.912428e+01 -1.910710e+01 -1.906438e+01 -1.902892e+01
## HQ(n)  -1.863398e+01 -1.856608e+01 -1.847264e+01 -1.838646e+01
## SC(n)  -1.790579e+01 -1.776255e+01 -1.759378e+01 -1.743227e+01
## FPE(n)  4.963245e-09  5.054624e-09  5.282030e-09  5.481070e-09
##                   9            10            11            12
## AIC(n) -1.901262e+01 -1.903088e+01 -1.902814e+01 -1.900245e+01
## HQ(n)  -1.831945e+01 -1.828698e+01 -1.823352e+01 -1.815711e+01
## SC(n)  -1.728993e+01 -1.718213e+01 -1.705334e+01 -1.690160e+01
## FPE(n)  5.581151e-09  5.491661e-09  5.519992e-09  5.679257e-09
```

```
mod_var <- VAR(train[c("Bankruptcy_Rate","Unemployment_Rate","House_Price_Index")],
           p=3, type="both", season=12, exogen = train["Population"])
fit_var <- data.frame(fitted(mod_var))
valid_var <- predict(mod_var, n.ahead=24, ci=0.95, dumvar = valid["Population"])
valid_pred <- data.frame(valid_var$fcst$Bankruptcy_Rate)
```

```
train_data %>%
  filter(!(Month %in% c(11987,21987,31987))) %>%
  ggplot()+
  geom_line(aes(x=Mon, y=Bankruptcy_Rate,color="Observed"))+
  geom_line(data=train[4:264,], aes(x=Mon, y=fit_var$Bankruptcy_Rate,color="Fitted"))+
  geom_line(data=valid, aes(x=Mon, y=valid_pred$fcst,color="Forecasted"))+
  labs(color='', x="Time", y="Bankruptcy Rate")
```



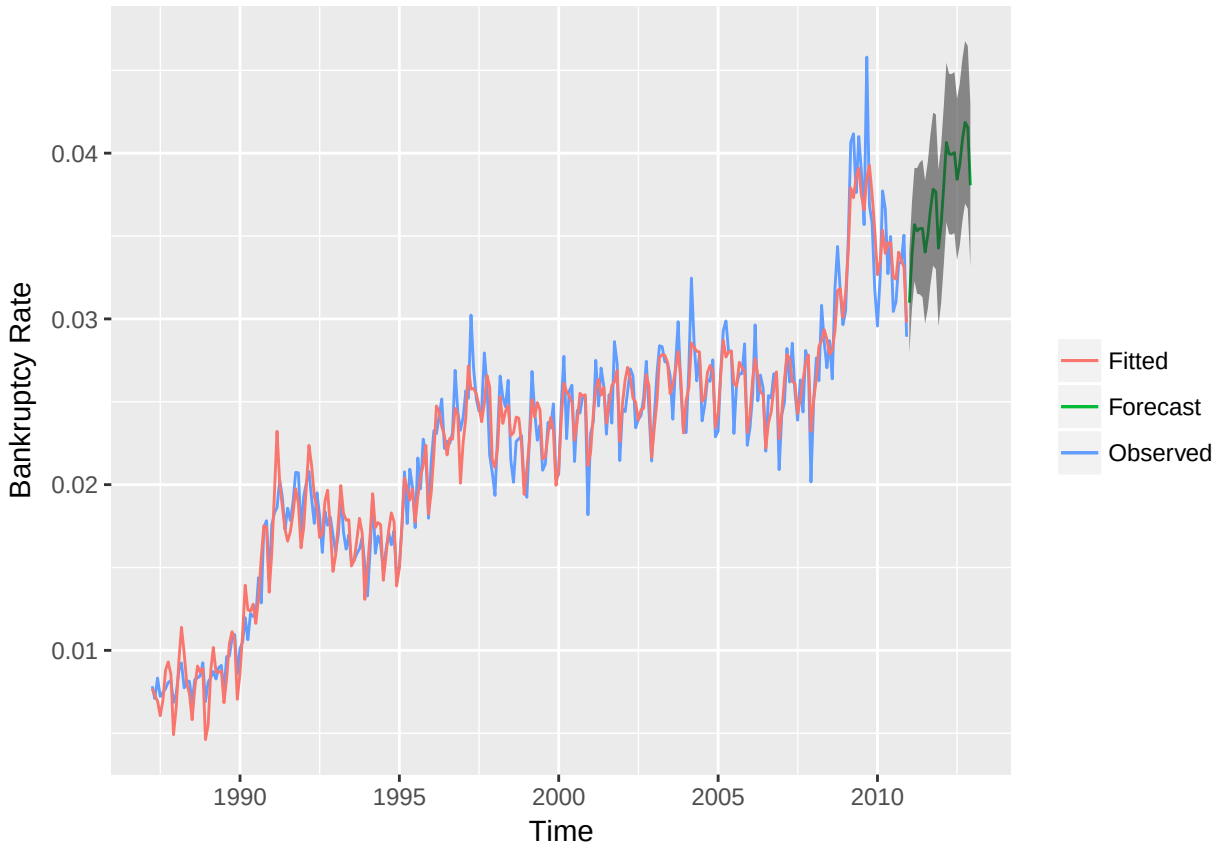

```
sqrt(mean((valid_pred$fcst - valid$Bankruptcy_Rate)^2))
```

```
## [1] 0.004338255
```

Retrain Model:

```
retrain_mod_var <- VAR(train_data[c("Bankruptcy_Rate","Unemployment_Rate","House_Price_Index")],
           p=3, type="both", season=12, exogen = train_data["Population"])
refit_var <- data.frame(fitted(retrain_mod_var))
test_pred <- data.frame(predict(retrain_mod_var, n.ahead=24, pi=0.95,
                                dumvar = test["Population"])$fcst$Bankruptcy_Rate)
```

```
train_data %>%
  filter(!(Month %in% c(11987,21987,31987))) %>%
  ggplot()+
  geom_line(aes(x=Mon, y=Bankruptcy_Rate, color='Observed'))+
  geom_line(aes(x=Mon, y=refit_var$Bankruptcy_Rate, color='Fitted'))+
  geom_line(data=test_pred, aes(x=test$Mon, y=fcst, color='Forecast'))+
  geom_ribbon(data=test_pred, aes(x=test$Mon, ymin=lower, ymax=upper, alpha=0.5), show.legend = F)+
  labs(color='',y='Bankruptcy Rate', x='Time')
```


Bankruptcy Rate
0.04
0.03
0.02
0.01
1990
1995
2000
2005
2010
Time
Fitted
Forecast
Observed