WEIGHTED AND COMBINED MODELS

Dr. Aric LaBarr

Institute for Advanced Analytics

COMPOSITE FORECASTING

Introduction

- The thought process around weighted combined forecasting (**also known as composite forecasting**) is **not new**.
- The topic has been studied for years, and **empirical evidence indicates that the combination of forecast methods tends to outperform most single forecast methods.**
- In 1969, James Bates and Clive Granger suggested that if the objective is to produce accurate forecasts, then the analyst should attempt to combine forecasts.

Combining Forecasts

- It is better to **average forecasts** in hope that by so doing, the biases among the methods and/or forecasters will compensate for one another.
- As a result, forecasts developed using different methods are expected to be more useful in cases of high uncertainty.
- This method is **especially relevant for long-range forecasting, where uncertainty is extremely high.**

Empirical Evidence

- Robert Winkler in 1967 and 1971 examined different weighting schemes in a study of the predication of football scores.
  - Found that the accuracy of a combined forecast was almost always as good as that of the best and most experienced forecaster.
- Additional studies have shown that the use of combined forecasts in areas such as finance, psychology, education, economics, and weather tend to improve overall forecast accuracy (Scott Armstrong 1985).

Empirical Evidence

- Bates and Granger (1969) Combined such methods as exponential smoothing and ARIMA
  - Overall, the errors of the combined forecasts were smaller than the error of either of the two components.
- Spyros Makridakis and Robert Winkler (1983) found that significant gains in accuracy were achieved as more forecasts were combined.
  - Also, combinations were more important for data measured on a shorter time interval (monthly) than on longer time intervals (years).

Two Basic Weighting Methods

1. Simple Averaging of Forecasts
2. Weighted Methods Based on Error
   i. Minimum Variance Weighting
   ii. Adaptive Weighting

SIMPLE AVERAGE

Two Basic Weighting Methods

1. Simple Averaging of Forecasts
2. Weighted Methods Based on Error
   i. Minimum Variance Weighting
   ii. Adaptive Weighting

Simple Averaging

$$\text{Combined Forecast} = \frac{1}{\text{n}} \sum_{i=1}^{n} FM_i$$

Simple Averaging

$$\text{Combined Forecast} = \frac{1}{n}\sum_{i=1}^{n} FM_i$$

Number of forecasting methods

Forecasting Method

Simple Averaging

```
For.Avg <- (HWES.USAir.train$mean +
            forecast::forecast(Full.ARIMA, xreg = cbind(Sep11, Sep11.L1, Sep11.L2, Sep11.L3,
                                                        Sep11.L4, Sep11.L5, Sep11.L6, Anniv), h = 12)$mean +
            Pass.Forecast +
            tail(predict(Prof, forecast.data)$yhat, 12))/4

Avg.error <- test - For.Avg

Avg.MAE <- mean(abs(Avg.error))
Avg.MAPE <- mean(abs(Avg.error)/abs(test))*100
```

Model Evaluation on Test Data

Model	MAE	MAPE
HW Exponential Smoothing	1134.58	1.76%
Seasonal ARIMA	1229.21	1.89%
Dynamic Regression ARIMA	1180.99	1.80%
Prophet	1449.85	2.25%
Neural Network AR	1087.85	1.67%
Avg of HW, ARIMA, Prophet, NNAR	1046.49	1.59%

Model Evaluation on Test Data

Model	MAE	MAPE
HW Exponential Smoothing	1134.58	1.76%
Seasonal ARIMA	1229.21	1.89%
Dynamic Regression ARIMA	1180.99	1.80%
Prophet	1449.85	2.25%
Neural Network AR	1087.85	1.67%
Avg of HW, ARIMA, Prophet, NNAR	1046.49	1.59%
Avg of HW, Prophet, NNAR	1001.66	1.52%

WEIGHTED AVERAGE

Two Basic Weighting Methods

1. Simple Averaging of Forecasts
2. Weighted Methods Based on Error
   1. Minimum Variance Weighting
   2. Adaptive Weighting

Minimum Variance Weighting

- Building composite forecasts require that the analyst select the weights to assign to the individual forecasts.
- Typically, we will assign weights that minimize the variance of the forecast errors over the forecasted period.

Minimum Variance – 2 Forecasts

$$\text{Combined Forecasts} = w_1 FM_1 + w_2 FM_2$$

Minimum Variance – 2 Forecasts

$$\text{Combined Forecasts} = w_1 FM_1 + w_2 FM_2$$

$$w_1 = \frac{\sigma^2{}_2 - \rho\sigma_1\sigma_2}{\sigma^2{}_1 + \sigma^2{}_2 - 2\rho\sigma_1\sigma_2}$$

$$w_2 = 1 - w_1$$

Minimum Variance – 2 Forecasts

$$\text{Combined Forecasts} = w_1 FM_1 + w_2 FM_2$$

$$w_1 = \frac{\sigma^2{}_2 - \rho\sigma_1\sigma_2}{\sigma^2{}_1 + \sigma^2{}_2 - 2\rho\sigma_1\sigma_2}$$

$$w_2 = 1 - w_1$$

For more than 2 forecasts, these equations get complicated.

Minimum Variance – Regression

$$Y_t = w_1 FM_1 + w_2 FM_2$$

- Run the regression of the actual values of Y against the two forecasted methods, **with the added restriction that $\boldsymbol{w_1 + w_2 = 1}$**.
- Get same results as previous slide.
- This makes it easy to extend to having more than 2 forecasts for Y:

$$Y_t = w_1 FM_1 + w_2 FM_2 + \cdots + w_k FM_k$$

$$w_1 + w_2 + \cdots + w_k = 1$$

Minimum Variance – Regression

```
Pass.Fit.NN <- rep(NA, 207)

for(i in 25:207){
  Pass.Fit.NN[i] <- training[i - 24] + NN.Model$fitted[i - 12]
}

Pass.Fit.ARIMA <- Full.ARIMA$fitted
Pass.Fit.HWES <- HWES.USAir.train$fitted
Pass.Fit.Prophet <- head(predict(Prof, forecast.data)$yhat, 207)

WC.Model <- lm(training ~ offset(Pass.Fit.HWES) +
                          I(Pass.Fit.ARIMA - Pass.Fit.HWES) +
                          I(Pass.Fit.NN - Pass.Fit.HWES) +
                          I(Pass.Fit.Prophet - Pass.Fit.HWES) - 1)

summary(WC.Model)
```

Minimum Variance – Regression

```
## Coefficients:
##                                       Estimate Std. Error t value Pr(>|t|)
## I(Pass.Fit.ARIMA - Pass.Fit.HWES)     0.74970    0.07001  10.714  < 2e-16 ***
## I(Pass.Fit.NN - Pass.Fit.HWES)       -0.02450    0.02774  -0.822    0.325
## I(Pass.Fit.Prophet - Pass.Fit.HWES)   0.28999    0.07306   3.956 9.79e-05 ***
```

```
ARIMA.coef <- coef(WC.Model)[1]
NN.coef <- coef(WC.Model)[2]
Prophet.coef <- coef(WC.Model)[3]
HW.coef <- 1 - ARIMA.coef - NN.coef - Prophet.coef
```

```
For.W.Avg <- HW.coef*HWES.USAir.train$mean +
             ARIMA.coef*forecast::forecast(Full.ARIMA, xreg = cbind(Sep11, Sep11.L1, Sep11.L2, Sep11.L3,
                                           Sep11.L4, Sep11.L5, Sep11.L6, Anniv), h = 12)$mean +
             NN.coef*Pass.Forecast +
             Prophet.coef*tail(predict(Prof, forecast.data)$yhat, 12)

W.Avg.error <- test - For.W.Avg

W.Avg.MAE <- mean(abs(W.Avg.error))
W.Avg.MAPE <- mean(abs(W.Avg.error)/abs(test))*100
```

Minimum Variance – Regression

```
## Coefficients:
##                                  Estimate Std. Error t value Pr(>|t|)
## I(Pass.Fit.ARIMA - Pass.Fit.HWES)    0.74970    0.07001  10.714  < 2e-16 ***
## I(Pass.Fit.NN - Pass.Fit.HWES)      -0.02450    0.02774  -0.822    0.325
## I(Pass.Fit.Prophet - Pass.Fit.HWES)  0.28999    0.07306   3.956 9.79e-05 ***

ARIMA.coef <- coef(WC.Model)[1]
NN.coef <- coef(WC.Model)[2]
Prophet.coef <- coef(WC.Model)[3]
HW.coef <- 1 - ARIMA.coef - NN.coef - Prophet.coef

For.W.Avg <- HW.coef*HWES.USAir.train$mean +
             ARIMA.coef*forecast::forecast(Full.ARIMA, xreg = cbind(Sep11, Sep11.L1, Sep11.L2, Sep11.L3,
                                           Sep11.L4, Sep11.L5, Sep11.L6, Anniv), h = 12)$mean +
             NN.coef*Pass.Forecast +
             Prophet.coef*tail(predict(Prof, forecast.data)$yhat, 12)

W.Avg.error <- test - For.W.Avg

W.Avg.MAE <- mean(abs(W.Avg.error))
W.Avg.MAPE <- mean(abs(W.Avg.error)/abs(test))*100
```

Model Evaluation on Test Data

Model	MAE	MAPE
HW Exponential Smoothing	1134.58	1.76%
Seasonal ARIMA	1229.21	1.89%
Dynamic Regression ARIMA	1180.99	1.80%
Prophet	1449.85	2.25%
Neural Network AR	1087.85	1.67%
Avg of HW, ARIMA, Prophet, NNAR	**1046.49**	**1.59%**
Avg of HW, Prophet, NNAR	**1001.66**	**1.52%**
Weighted Avg of HW, ARIMA, Prophet, NNAR	1068.44	1.61%

Two Basic Weighting Methods

1. Simple Averaging of Forecasts
2. Weighted Methods Based on Error
   i. Minimum Variance Weighting
   ii. Adaptive Weighting

Adaptive Weighting

- Adaptive weighting is structured in much the same way as minimum variance weighting.
- The only difference is that we will weight more recent time periods more heavily.
  - If forecast method 1 does great for the first half of the data, but forecast method 2 does better for the second half, wouldn't you prefer to have forecast method 2 since it does better on recent data?