

Rob J Hyndman  
George Athanasopoulos

# FORECASTING

## PRINCIPLES AND PRACTICE

A comprehensive introduction to the latest forecasting methods using R. Learn to improve your forecast accuracy using dozens of real data examples.



3RD EDITION

 **OTexts**  
OPEN TEXTS FOR PRACTICE

## 7. Time series regression models

### 7.1 The linear model

[OTexts.org/fpp3/](https://OTexts.org/fpp3/)

# Least squares estimation

- In practice we need to estimate the coefficients:  $\beta_0, \beta_1, \dots, \beta_k$ .

$$\sum_{t=1}^T \varepsilon_t^2 = \sum_{t=1}^T (y_t - \beta_0 - \beta_1 x_{1,t} - \beta_2 x_{2,t} - \dots - \beta_k x_{k,t})^2$$

# Least squares estimation

- In practice we need to estimate the coefficients:  $\beta_0, \beta_1, \dots, \beta_k$ .

$$\sum_{t=1}^T \varepsilon_t^2 = \sum_{t=1}^T (y_t - \beta_0 - \beta_1 x_{1,t} - \beta_2 x_{2,t} - \dots - \beta_k x_{k,t})^2$$

```
model(TSLM(y ~ x_1 + x_2 + ... + x_k))
```

- Estimated coefficients:  $\hat{\beta}_0, \dots, \hat{\beta}_k$

# Example: US consumption expenditure

```
fit_consMR <- us_change |>
  model(lm = TSLM(Consumption ~ Income + Production + Unemployment + Savings))
report(fit_consMR)
```

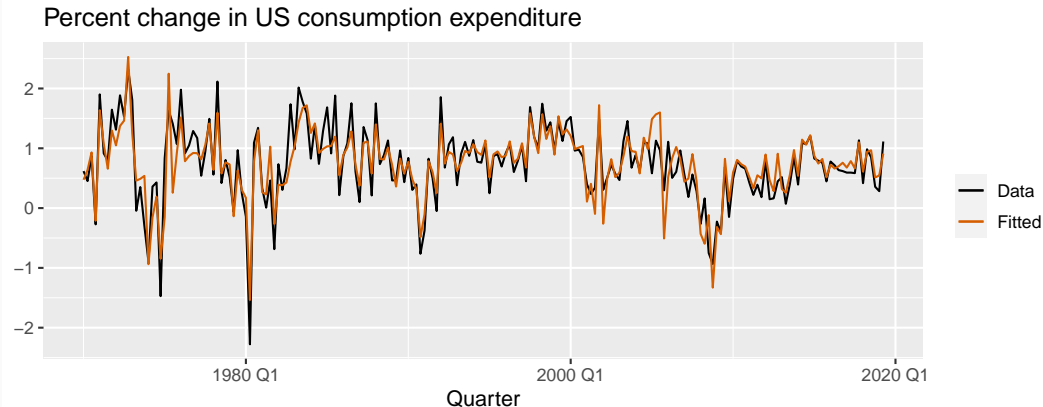
```
## Series: Consumption
## Model: TSLM
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.906 -0.158 -0.036  0.136  1.155
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   0.25311    0.03447   7.34 5.7e-12 ***
## Income        0.74058    0.04012  18.46 < 2e-16 ***
## Production    0.04717    0.02314   2.04  0.043 *
## Unemployment -0.17469    0.09551  -1.83  0.069 .
## Savings       -0.05289    0.00292 -18.09 < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.31 on 193 degrees of freedom
## Multiple R-squared:  0.768,    Adjusted R-squared:  0.763
## F-statistic: 160 on 4 and 193 DF, p-value: <2e-16
```

# Fitted values

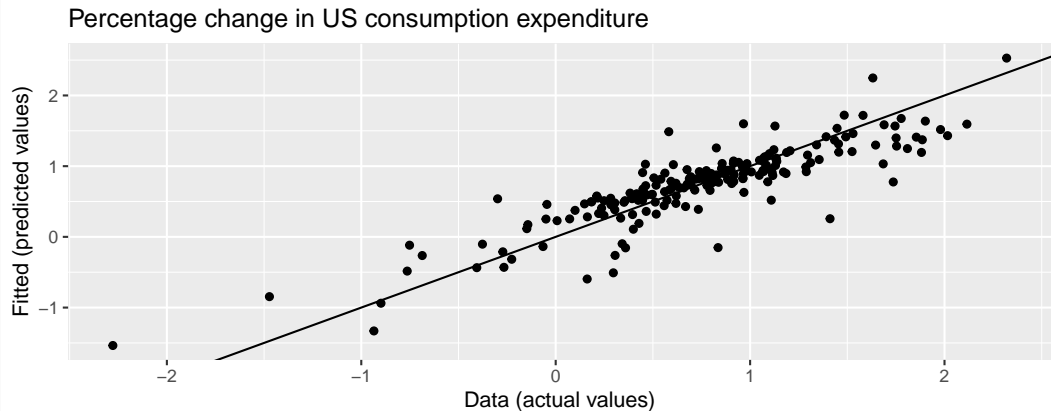
$$\hat{y}_t = \hat{\beta}_0 + \hat{\beta}_1 x_{1,t} + \hat{\beta}_2 x_{2,t} + \cdots + \hat{\beta}_k x_{k,t}$$

# Fitted values

$$\hat{y}_t = \hat{\beta}_0 + \hat{\beta}_1 x_{1,t} + \hat{\beta}_2 x_{2,t} + \cdots + \hat{\beta}_k x_{k,t}$$



# Example: US consumption expenditure



# Goodness of fit

## Coefficient of determination

$$R^2 = \frac{\sum(\hat{y}_t - \bar{y})^2}{\sum(y_t - \bar{y})^2}$$

## Standard error of the regression

$$\hat{\sigma}_e = \sqrt{\frac{1}{T - k - 1} \sum_{t=1}^T e_t^2}$$

where  $k$  is the number of predictors in the model.