Seasonal Models

Dr. Jarad Niemi

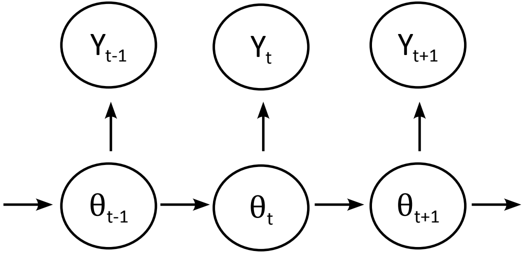
STAT 6150 - Iowa State University

October 9, 2025

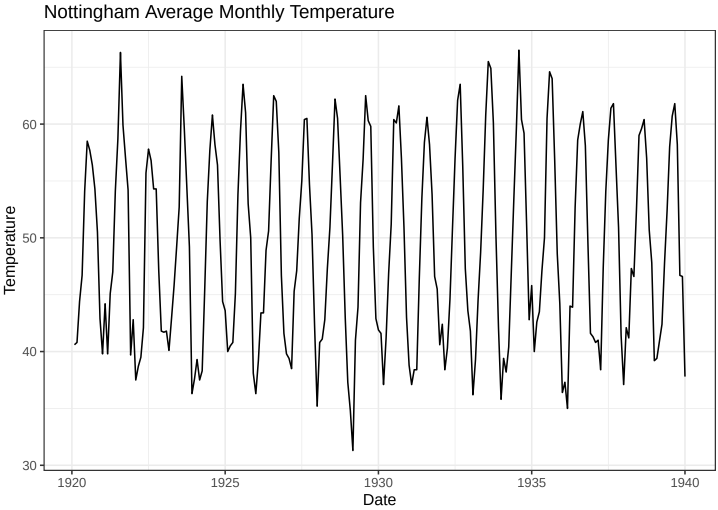
Dynamic Linear Models

$$
\begin{aligned}
Y_t &= F_t\theta_t + v_t & v_t &\stackrel{ind}{\sim} N_m(0, V_t) \\
\theta_t &= G_t\theta_{t-1} + w_t & w_t &\stackrel{ind}{\sim} N_p(0, W_t) \\
& & \theta_0 &\sim N_p(m_0, C_0)
\end{aligned}
$$

where v_t and w_t are independent across time and all are independent of θ_0.



Nottem Data


Nottingham Average Monthly Temperature
Temperature
60
50
40
30
1920
1925
1930
1935
1940
Date

Regress - air temp on month

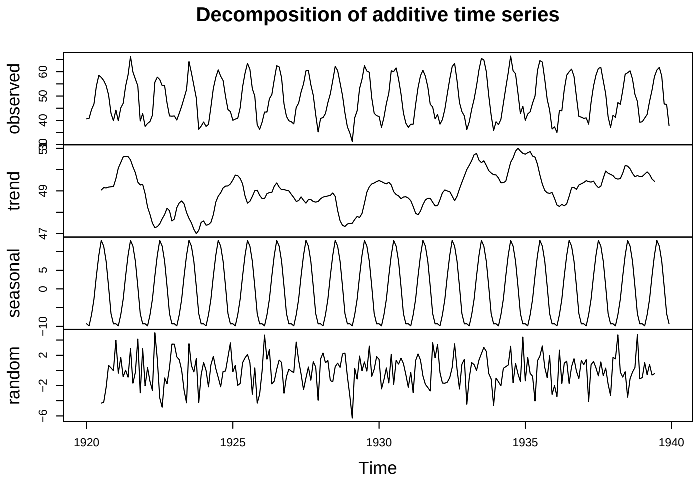
```
##
## Call:
## lm(formula = Temperature ~ as.factor(Month), data = nottem_d)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -7.890 -1.369  0.285  1.405  6.270
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         39.6950     0.5176  76.691  < 2e-16 ***
## as.factor(Month)2   -0.5050     0.7320  -0.690 0.490957
## as.factor(Month)3    2.5000     0.7320   3.415 0.000754 ***
## as.factor(Month)4    6.5950     0.7320   9.010  < 2e-16 ***
## as.factor(Month)5   12.8650     0.7320  17.575  < 2e-16 ***
## as.factor(Month)6   18.3450     0.7320  25.062  < 2e-16 ***
## as.factor(Month)7   22.2050     0.7320  30.335  < 2e-16 ***
## as.factor(Month)8   20.8250     0.7320  28.450  < 2e-16 ***
## as.factor(Month)9   16.7850     0.7320  22.931  < 2e-16 ***
## as.factor(Month)10   9.8000     0.7320  13.388  < 2e-16 ***
## as.factor(Month)11   2.8850     0.7320   3.941 0.000108 ***
## as.factor(Month)12  -0.1650     0.7320  -0.225 0.821859
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.315 on 228 degrees of freedom
## Multiple R-squared:  0.9304,Adjusted R-squared:  0.9271
## F-statistic: 277.3 on 11 and 228 DF,  p-value: < 2.2e-16
```

Estimated Monthly Means

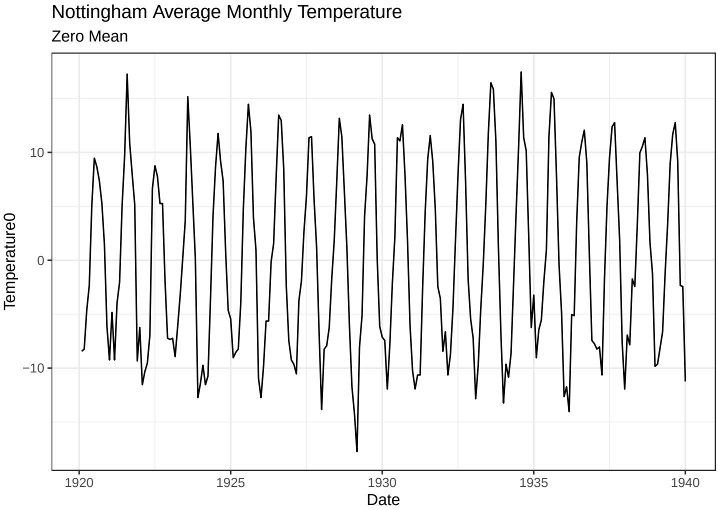
```
em <- emmeans(m, ~ Month); em

##  Month emmean    SE  df lower.CL upper.CL
##      1   39.7 0.518 228     38.7     40.7
##      2   39.2 0.518 228     38.2     40.2
##      3   42.2 0.518 228     41.2     43.2
##      4   46.3 0.518 228     45.3     47.3
##      5   52.6 0.518 228     51.5     53.6
##      6   58.0 0.518 228     57.0     59.1
##      7   61.9 0.518 228     60.9     62.9
##      8   60.5 0.518 228     59.5     61.5
##      9   56.5 0.518 228     55.5     57.5
##     10   49.5 0.518 228     48.5     50.5
##     11   42.6 0.518 228     41.6     43.6
##     12   39.5 0.518 228     38.5     40.5
##
## Confidence level used: 0.95
```

Time Series Decomposition



Nottem Data - Zero Mean



Seasonal factors

Suppose all s seasons have a factor α_i. Then

$$
\begin{array}{rcl}
Y_1 & = & \alpha_1 + v_1 \\
Y_2 & = & \alpha_2 + v_2 \\
& \vdots & \\
Y_s & = & \alpha_s + v_s \\
Y_{s+1} & = & \alpha_1 + v_{s+1} \\
Y_{s+2} & = & \alpha_2 + v_{s+2} \\
& \vdots & \\
Y_{2s} & = & \alpha_s + v_{2s} \\
Y_{2s+1} & = & \alpha_1 + v_{2s+1} \\
Y_{2s+2} & = & \alpha_2 + v_{2s+2} \\
& \vdots &
\end{array}
$$

$$
\begin{array}{rcl}
Y_t & = & F_t\theta_t + v_t \\
\theta_t & = & G_t\theta_{t-1} + w_t
\end{array}
$$

where

$$
\begin{array}{rcl}
\theta_1 & = & (\alpha_1, \alpha_2, \ldots, \alpha_s)^\top \\
F_t & = & F = (1, 0, 0, \ldots, 0) \\
G_t & = & G = \left[\begin{array}{ccccc} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ 1 & 0 & 0 & \vdots & 0 \end{array}\right] \\
W_t & = & W = 0 \quad \text{(maybe)}
\end{array}
$$

Rotating factors

If $\theta_1 = (\alpha_1, \alpha_2, \ldots, \alpha_s)^\top$ and $W_t = 0$, what is θ_t?

$$\begin{aligned}
\theta_2 &= G\theta_1 = (\alpha_2, \alpha_3, \alpha_4, \ldots, \alpha_s, \alpha_1)^\top \\
\theta_3 &= G\theta_2 = (\alpha_3, \alpha_4, \ldots, \alpha_s, \alpha_1, \alpha_2)^\top \\
&\vdots \\
\theta_t &= G\theta_{t-1} = (\alpha_j, \alpha_{j+1}, \ldots, \alpha_s, \alpha_1, \ldots, \alpha_{j-1})^\top
\end{aligned}$$

where $j = t \bmod s$ ($j\%\%s$ in R).

Alternative DLM for seasonal factors

Suppose all s seasons have a factor α_i. Then

$$
\begin{array}{rcl}
Y_1 &=& \alpha_1 + v_1 \\
Y_2 &=& \alpha_2 + v_2 \\
&\vdots& \\
Y_s &=& \alpha_s + v_s \\
Y_{s+1} &=& \alpha_1 + v_{s+1} \\
Y_{s+2} &=& \alpha_2 + v_{s+2} \\
&\vdots& \\
Y_{2s} &=& \alpha_s + v_{2s} \\
Y_{2s+1} &=& \alpha_1 + v_{2s+1} \\
Y_{2s+2} &=& \alpha_2 + v_{2s+2} \\
&\vdots&
\end{array}
$$

$$
\begin{array}{rcl}
Y_t &=& F_t\theta_t + v_t \\
\theta_t &=& G_t\theta_{t-1} + w_t
\end{array}
$$

where

$$
\begin{array}{rcl}
\theta_1 &=& (\alpha_1, \alpha_s, \alpha_{s-1}, \ldots, \alpha_2)^\top \\
F_t &=& F = (1, 0, 0, \ldots, 0) \\
G_t &=& G = \left[\begin{array}{ccccc}
0 & 0 & \cdots & 0 & 1 \\
1 & 0 & \cdots & 0 & 0 \\
0 & 1 & \cdots & 0 & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & 0 & 0 \\
0 & 0 & \vdots & 1 & 0
\end{array}\right] \\
W_t &=& W = 0 \quad \text{(maybe)}
\end{array}
$$

Identifiability in seasonal factor models

- Modeling mean separately
- Looking at deviations from the mean

$\Longrightarrow$ parameter identifiability issue

- Identifiability constraints:
  - Set $\alpha_j = 0$ for some $j \in \{1, 2, \dots, s\}$.
  - Sum-to-zero constraint, i.e. $\sum_{i=1}^{s} \alpha_i = 0$.

Parsimonious seasonal factor model

$$\begin{array}{rcl} Y_t &=& F_t\theta_t + v_t \\ \theta_t &=& G_t\theta_{t-1} + w_t \end{array}$$

where

$$\begin{array}{rcl} \theta_1 &=& (\alpha_1, \alpha_s, \alpha_{s-1}, \ldots, \alpha_3)^\top \\ F_t &=& F = (1, 0, \ldots, 0) \\ G_t &=& G = \left[\begin{array}{ccccc} -1 & -1 & \cdots & -1 & -1 \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & 0 \\ & & \vdots & & \\ 0 & 0 & \vdots & 1 & 0 \end{array}\right] \\ W_t &=& W = 0 \quad \text{(maybe)} \end{array}$$

What is θ_2 if $W = 0$?

$$\begin{array}{rcl} \theta_2 &=& G\theta_1 = (\alpha_2, \alpha_1, \alpha_s, \alpha_{s-1}, \ldots, \alpha_4)^\top \\ \theta_3 &=& G\theta_2 = (\alpha_3, \alpha_2, \alpha_1, \alpha_s, \alpha_{s-1}, \ldots, \alpha_5)^\top \\ & \vdots & \\ \theta_t &=& G\theta_{t-1} \\ &=& (\alpha_j, \alpha_{j-1}, \ldots, \alpha_1, \alpha_s, \alpha_{s-1}, \ldots, \alpha_{j+2})^\top \end{array}$$

where $j = t \bmod s$.

Seasonal factor model in R

```
build_sf_model <- function(theta, ...) {
  dlmModSeas(frequency = 12,
             dV = exp(theta),
             dW = rep(0,11),  # seasonality is constant
             ...)
}

nottem_MLEs <- dlmMLE(nottem_d$Temperature0, 0, build_sf_model)
stopifnot(!nottem_MLEs$convergence) # 0 is converged
exp(nottem_MLEs$par) # estimated V
```

```
## [1] 5.334667
```

```
# Fit seasonal factor model
nottem_sf_model <- dlmModSeas(12, dV = exp(nottem_MLEs$par), dW = rep(0,11))
nottemFilt      <- dlmFilter(nottem_d$Temperature0, nottem_sf_model)

# Extract filtered distribution means at time T
m_T <- c(-sum(nottemFilt$m[241,]), rev(nottemFilt$m[241,]))
```

Compare estimates to lm

```
m0 <- lm(Temperature0 ~ as.factor(Month) + 0, data = nottem_d)
cbind("lm" = coef(m0),
      "dlm" = m_T)
```

```
##                             lm        dlm
## as.factor(Month)1  -9.3445833 -9.3445833
## as.factor(Month)2  -9.8495833 -9.8495830
## as.factor(Month)3  -6.8445833 -6.8445831
## as.factor(Month)4  -2.7495833 -2.7495832
## as.factor(Month)5   3.5204167  3.5204166
## as.factor(Month)6   9.0004167  9.0004164
## as.factor(Month)7  12.8604167 12.8604163
## as.factor(Month)8  11.4804167 11.4804164
## as.factor(Month)9   7.4404167  7.4404165
## as.factor(Month)10  0.4554167  0.4554167
## as.factor(Month)11 -6.4595833 -6.4595831
## as.factor(Month)12 -9.5095833 -9.5095831
```

Canonical basis

Suppose all s seasons have a factor α_i. Then, think of α as a linear combination of basis vectors

$$\alpha = (\alpha_1, \ldots, \alpha_s) = \sum_{i=1}^{s} \alpha_i u_i$$

where the u_i form the canonical basis, i.e. u_i is the s-dimensional vector having the ith component equal to 1 and all other elements 0

- This representation lacks
  - Smoothness differentiation
  - Parsimony
  - Interpretation (perhaps)

Fourier frequencies

Let

$$\omega_j = 2\pi \frac{j}{s}, \quad j = 0, 1, \ldots, \frac{s}{2}$$

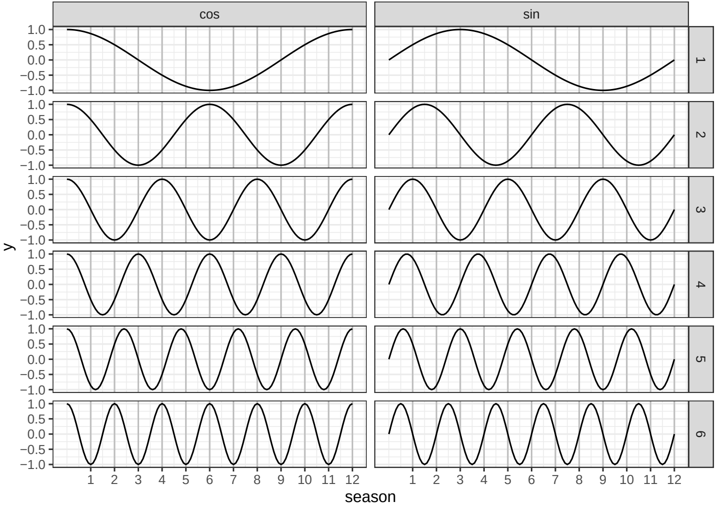
where s is the length of the period.

Consider

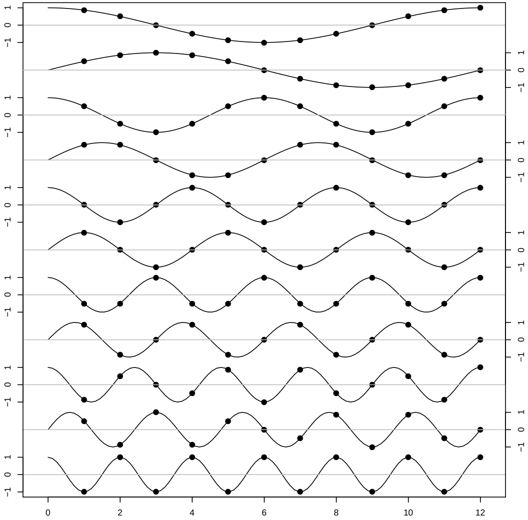
$$\begin{aligned}
e_0 &= (1, 1, \ldots, 1)^\top \\
c_1 &= (\cos\omega_1, \cos 2\omega_1, \ldots \cos s\omega_1)^\top \\
s_1 &= (\sin\omega_1, \sin 2\omega_1, \ldots \sin s\omega_1)^\top \\
&\vdots \\
c_j &= (\cos\omega_j, \cos 2\omega_j, \ldots \cos s\omega_j)^\top \\
s_j &= (\sin\omega_j, \sin 2\omega_j, \ldots \sin s\omega_j)^\top \\
&\vdots \\
c_{s/2} &= (\cos\omega_{s/2}, \cos 2\omega_{s/2}, \ldots \cos s\omega_{s/2})^\top
\end{aligned}$$

$s_{s/2}$ is all zeros.

Trigonometric basis functions



Plotting Fourier basis



Fourier basis

Basis for $\mathbb{R}^s$

$$\alpha = a_0 e_0 + \sum_{j=1}^{s/2-1} (a_j c_j + b_j s_j) + a_{s/2} c_{s/2}.$$

Assume $a_0 = 0$ since the mean will be modeled separately, e.g. through a polynomial trend model. For $j = 1, 2, \ldots, s/2$, the jth harmonic is

$$\begin{aligned} S_j(t) &= a_j \cos(t\omega_j) + b_j \sin(t\omega_j) \\ &= A_j \cos(t\omega_j + \gamma_j) \end{aligned}$$

where $b_{s/2} = 0$, $A_j = \sqrt{a_j^2 + b_j^2}$ is the amplitude, and $\gamma_j = \text{arctan}(-b_j/a_j)$ is the phase.

Evolving a harmonic

For $j = 1, 2, \ldots, s/2$, the jth harmonic is

$$S_j(t) = a_j \cos(t\omega_j) + b_j \sin(t\omega_j)$$

What is $S_j(t+1)$? It is

$$S_j(t+1) = a_j \cos([t+1]\omega_j) + b_j \sin([t+1]\omega_j).$$

For $j < s/2$, if we only know the value of $S_j(t)$, i.e. we don't know the value of a_j and b_j individual, we cannot determine $S_j(t+1)$. But we can find that

$$\begin{aligned}
S_j(t+1) = \;& a_j \cos([t+1]\omega_j) + b_j \sin([t+1]\omega_j) \\
& \vdots \\
= \;& (a_j \cos(t\omega_j) + b_j \sin(t\omega_j)) \cos(\omega_j) + \\
& + (-a_j \sin(t\omega_j) + b_j \cos(t\omega_j)) \sin(\omega_j) \\
\\
= \;& S_j(t) \cos(\omega_j) + S_j^*(t) \sin(\omega_j)
\end{aligned}$$

Building a DLM - constructing G

For $j = 1, 2, \ldots, s/2$, the jth harmonic is

$$S_j(t) = a_j \cos(t\omega_j) + b_j \sin(t\omega_j)$$

The conjugate harmonic is

$$S_j^*(t) = -a_j \sin(t\omega_j) + b_j \cos(t\omega_j).$$

And the jth harmonic evolves in time according to

$$\left[\begin{array}{c} S_j(t+1) \\ S_j^*(t+1) \end{array} \right] = \left[\begin{array}{cc} \cos\omega_j & \sin\omega_j \\ -\sin\omega_j & \cos\omega_j \end{array} \right] \left[\begin{array}{c} S_j(t) \\ S_j^*(t) \end{array} \right]$$

and

$$S_{s/2}(t+1) = -S_{s/2}(t)$$

Period j DLMs

Consider DLMs containing only seasonality of frequency $j < s/2$.

$$\begin{bmatrix} S_j(t+1) \\ S_j^*(t+1) \end{bmatrix} = \begin{bmatrix} \cos\omega_j & \sin\omega_j \\ -\sin\omega_j & \cos\omega_j \end{bmatrix} \begin{bmatrix} S_j(t) \\ S_j^*(t) \end{bmatrix}.$$

So this is a DLM with evolution matrix

$$H_j = \begin{bmatrix} \cos\omega_j & \sin\omega_j \\ -\sin\omega_j & \cos\omega_j \end{bmatrix},$$

state vector $(S_j(t), S_j^*(t))^\top$, and observation matrix $F = [1\,0]$.

If $j = s/2$, we have

$$S_{s/2}(t+1) = \cos((t+1)\pi) = -\cos(t\pi) = -S_{s/2}(t)$$

which is a DLM with state vector $S_{s/2}(t)$, evolution matrix $H_{s/2} = [-1]$, and observation matrix $F = [1]$.

Fourier form seasonal DLM

Combine these period j DLMs using our combining rules:

$$\theta_t = (S_1(t), S_1^*(t), \ldots, S_{\frac{s}{2}-1}(t), S^*_{\frac{s}{2}-1}(t), S_{\frac{s}{2}}(t))^\top$$

with evolution matrix

$$G = \text{blockdiag}(H_1, \ldots, H_{\frac{s}{2}})$$

and observation matrix

$$F = [1\,0\,1\,0\,\ldots\,0\,1]$$

and initial vector

$$\theta_0 = (a_1, b_1, \ldots, a_{\frac{s}{2}-1}, b_{\frac{s}{2}-1}, a_{\frac{s}{2}})^\top.$$

The seasonal process can be made more smooth by dropping higher harmonics.