Fourier Features (part 1)

Seasonality features

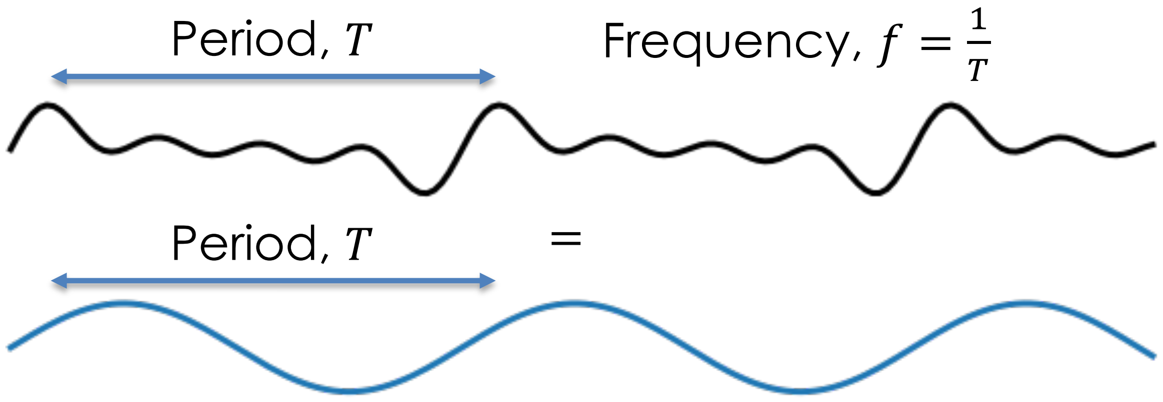
What is a Fourier series?

Period, T

Frequency, $f = \frac{1}{T}$

$$s(t) = s(t + T)$$

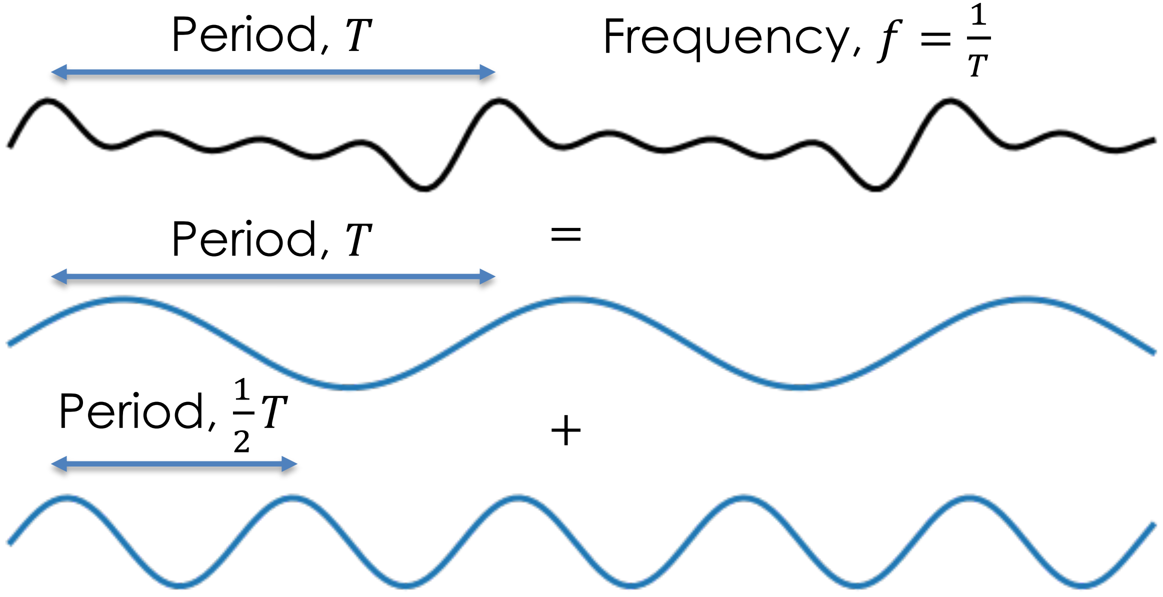
What is a Fourier series?



$s(t) = s(t + T)$

$A_1 \sin(2\pi * f * t)$

What is a Fourier series?

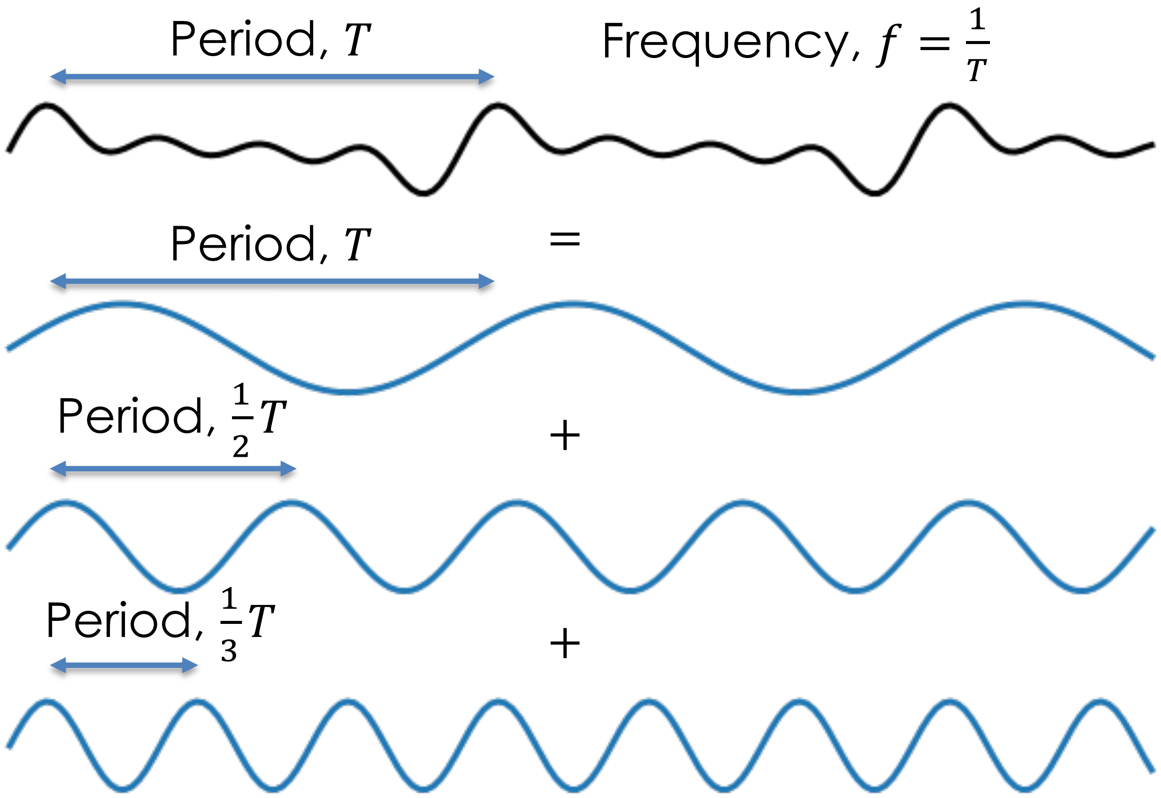


$s(t) = s(t + T)$

$A_1 \sin(2\pi * f * t)$

$A_2 \sin(2\pi * 2f * t)$

What is a Fourier series?



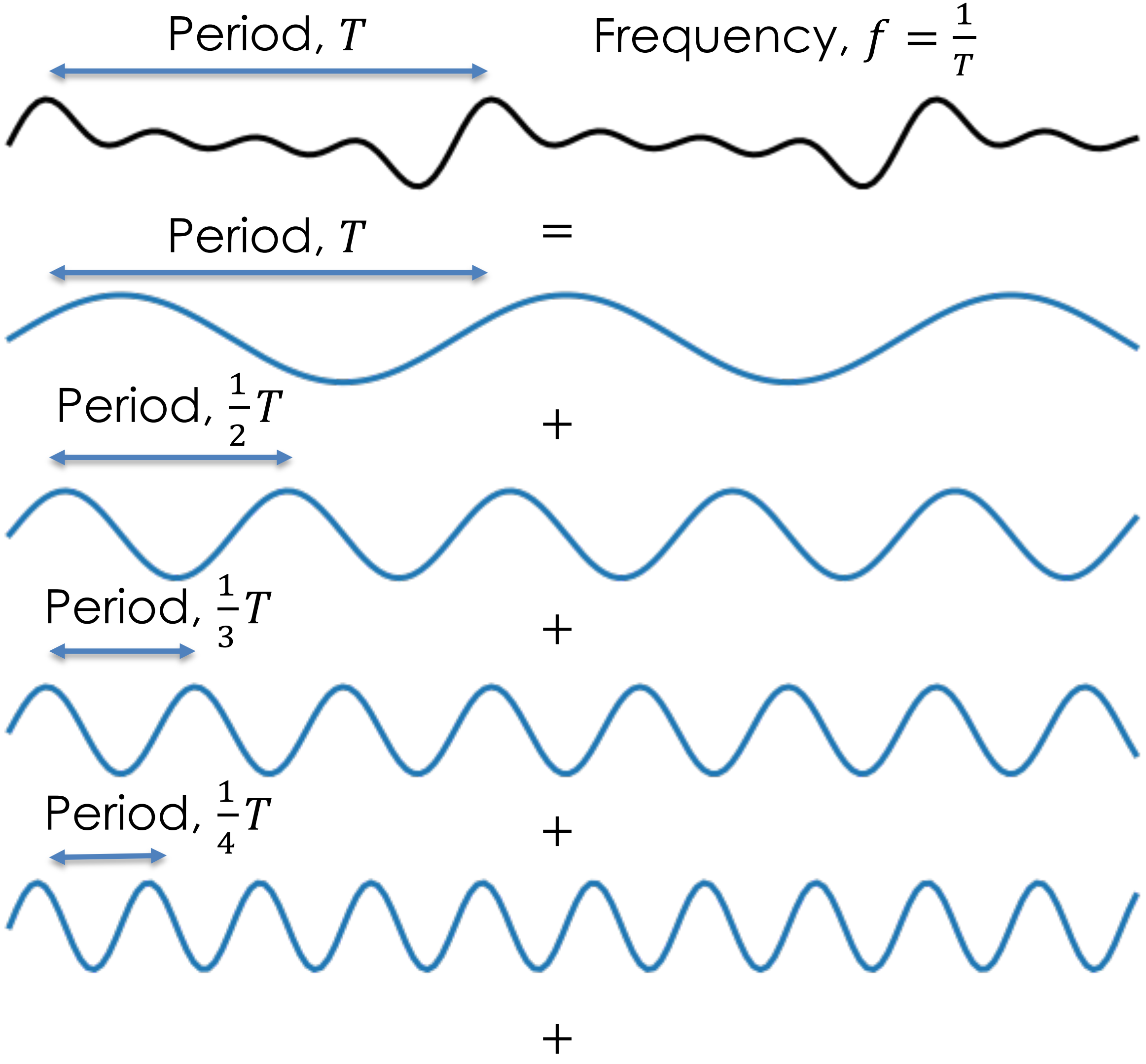
$s(t) = s(t + T)$

$A_1 \sin(2\pi * f * t)$

$A_2 \sin(2\pi * 2f * t)$

$A_3 \sin(2\pi * 3f * t)$

What is a Fourier series?



$s(t) = s(t + T)$

$A_1 \sin(2\pi * f * t)$

$A_2 \sin(2\pi * 2f * t)$

$A_3 \sin(2\pi * 3f * t)$

$A_4 \sin(2\pi * 4f * t)$

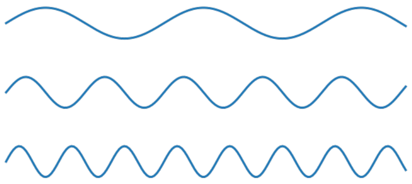
What is a Fourier series?

Period, T Frequency, $f = \frac{1}{T}$

$$s(t) = s(t+T)$$

- Any periodic function, $s(t)$, can be written as a Fourier series expansion:

$$\begin{aligned} s(t) = A_0 &+ A_1 \sin(2\pi * f * t) + B_1 \cos(2\pi * f * t) \\ &+ A_2 \sin(2\pi * 2f * t) + B_2 \cos(2\pi * 2f * t) \\ &+ A_3 \sin(2\pi * 3f * t) + B_3 \cos(2\pi * 3f * t) \\ &+ \qquad \cdots \end{aligned}$$

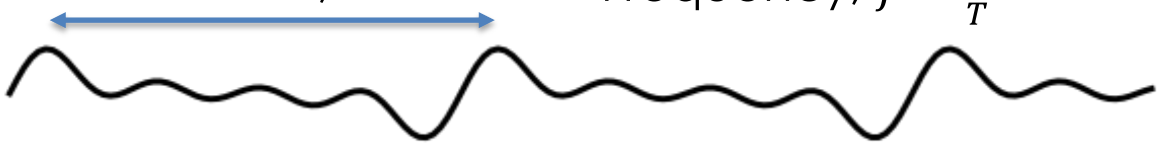


…

What is a Fourier series?

Period, T

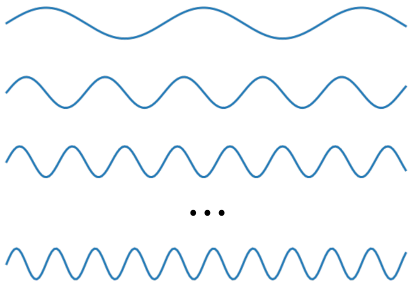
Frequency, $f = \frac{1}{T}$



$$s(t) = s(t + T)$$

- Any periodic function, $s(t)$, can be written as a Fourier series expansion:

$$\begin{aligned} s(t) \approx A_0 &+ A_1 \sin(2\pi * f * t) + B_1 \cos(2\pi * f * t) \\ &+ A_2 \sin(2\pi * 2f * t) + B_2 \cos(2\pi * 2f * t) \\ &+ A_3 \sin(2\pi * 3f * t) + B_3 \cos(2\pi * 3f * t) \\ &+ \qquad \cdots \\ &+ A_N \sin(2\pi * Nf * t) + B_N \cos(2\pi * Nf * t) \end{aligned}$$



- This allows us to approximate **any** periodic function (e.g., seasonality component!) as a sum of sines and cosines!

How is this related to forecasting?

- Additive time series decomposition:

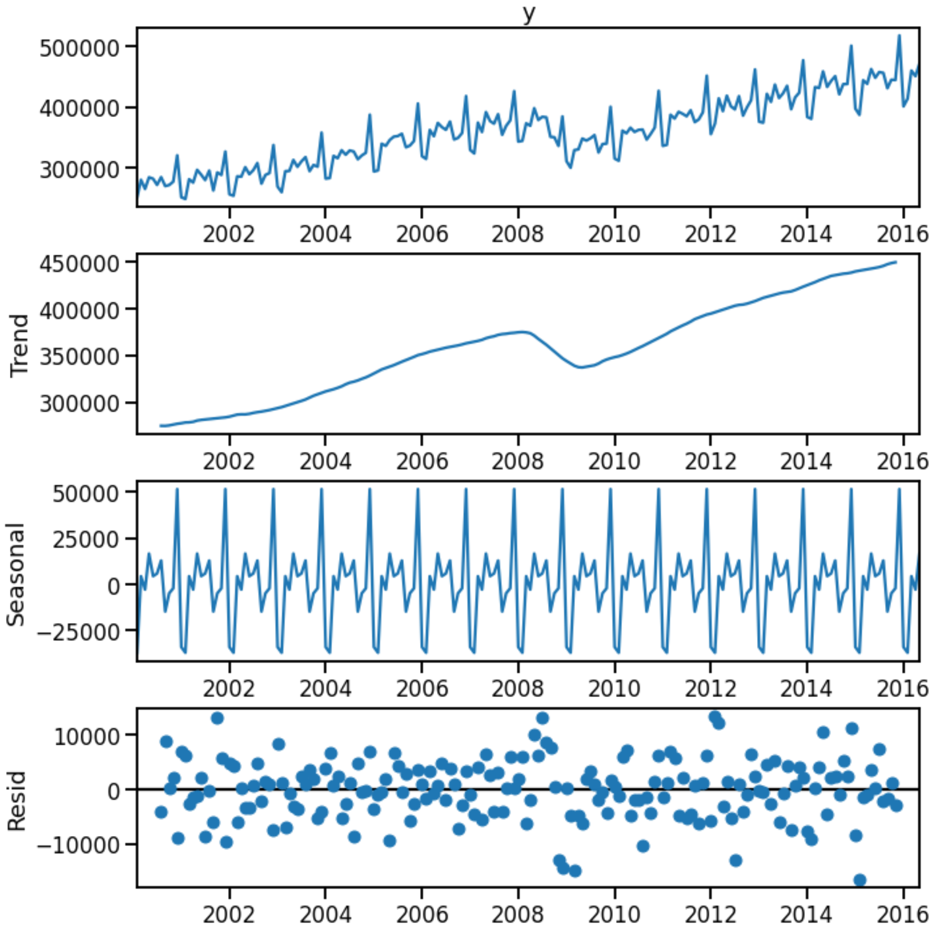
$$y_t = trend_t + seasonality_t + residuals_t$$

- Linear model:

$$y_t = \beta_0 \ + \beta_1 x_1 + \beta_2 x_2 + \cdots$$

- If we think of $seasonality_t$ as a periodic function with frequency f we can represent it using a Fourier series:

$$seasonality_t \approx A_0 + \sum_{n=1}^{N} A_n \sin(2\pi n f t) + B_n \cos(2\pi n f t)$$



How is this related to forecasting?

- Additive time series decomposition:

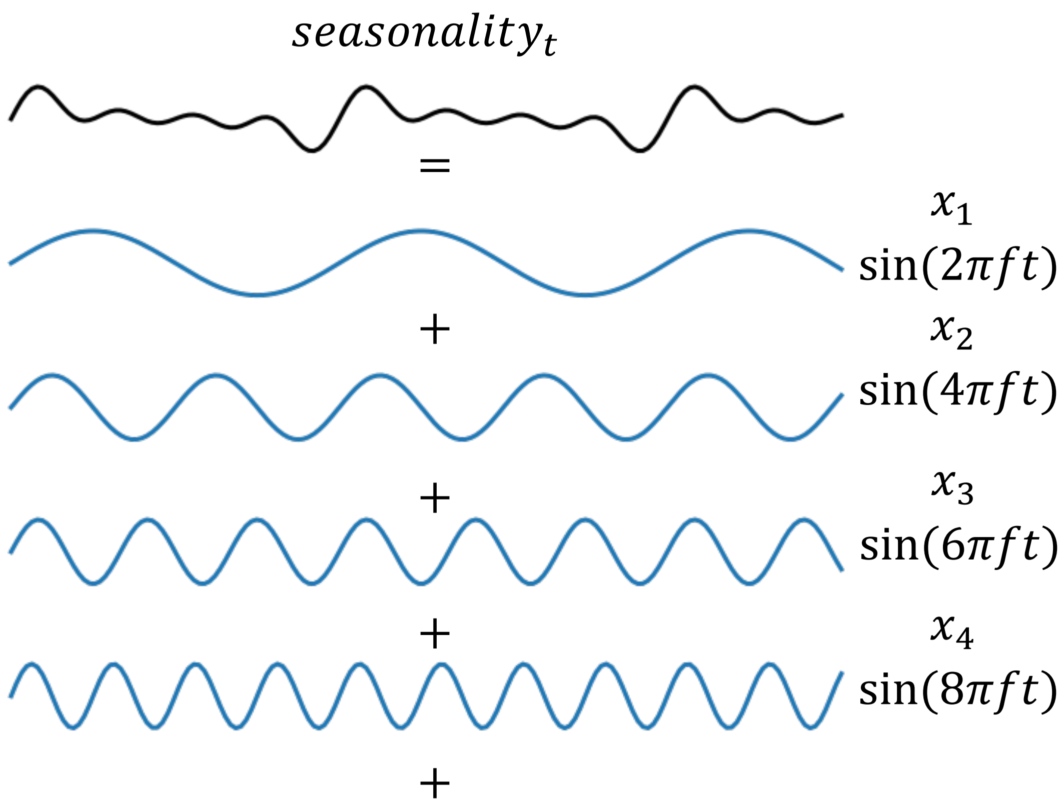
$$y_t = trend_t + seasonality_t + residuals_t$$

- Linear model:

$$y_t = \beta_0 \ + \beta_1 x_1 + \beta_2 x_2 + \cdots$$

- If we think of $seasonality_t$ as a periodic function with frequency f we can represent it using a Fourier series:

$$seasonality_t \approx A_0 + \sum_{n=1}^{N} A_n \sin(2\pi n f t) + B_n \cos(2\pi n f t)$$



How is this related to forecasting?

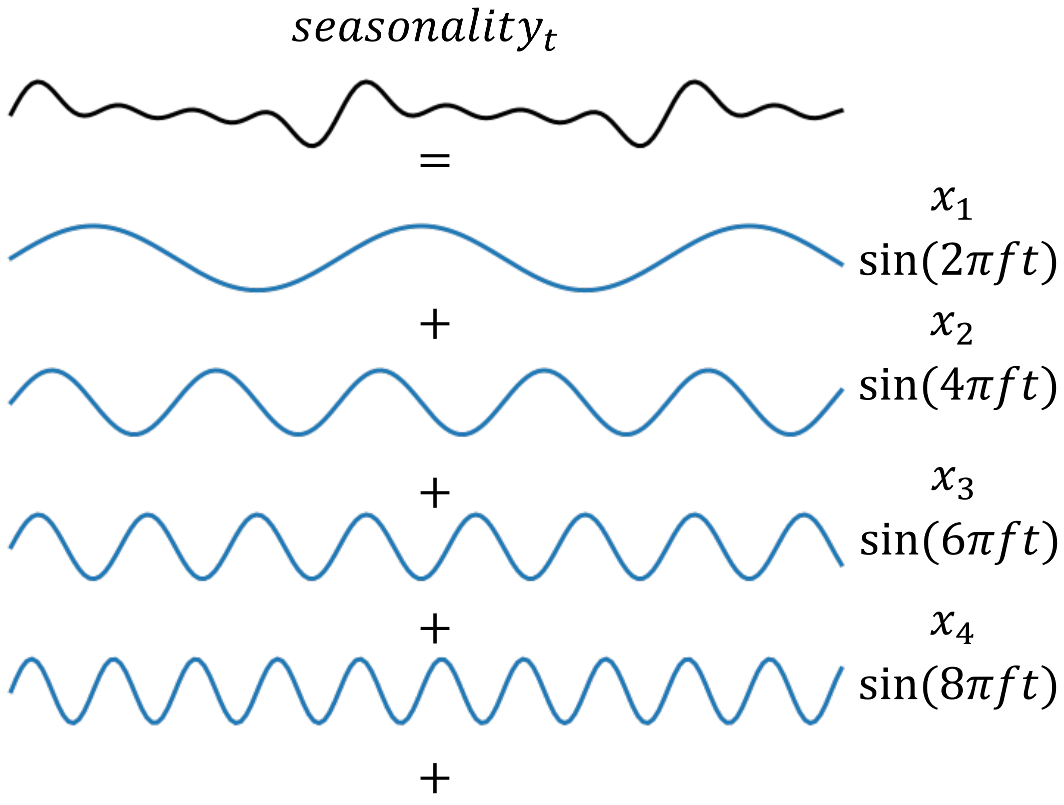
- We can use Fourier terms (sine & cosines) as features in a linear model to achieve this.

- Linear model:

$$
\begin{aligned}
y_t &= \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots \\
&= \beta_0 + \beta_1 \sin(2\pi f t) + \beta_2 \sin(4\pi f t) + \ \ldots
\end{aligned}
$$

- The Fourier coefficients, A_n and B_n, then become coefficients in the linear model, β_n, that are learned.

- The number of Fourier terms, N, is a hyperparameter controlling how complex the fitted seasonality is.
  - N is too large: overfit to noise.
  - N is too small: underfit the true seasonal pattern.



How is this related to forecasting?

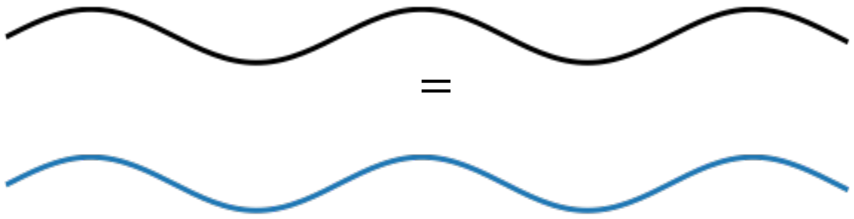
- We can use Fourier terms (sine & cosines) as features in a linear model to achieve this.

- Linear model:

$$
\begin{aligned}
y_t &= \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots \\
&= \beta_0 + \beta_1 \sin(2\pi f t) + \beta_2 \sin(4\pi f t) + \ldots
\end{aligned}
$$

- The Fourier coefficients, A_n and B_n, then become coefficients in the linear model, β_n, that are learned.

- The number of Fourier terms, N, is a hyperparameter controlling how complex the fitted seasonality is.
  - N is too large: overfit to noise.
  - N is too small: underfit the true seasonal pattern.



How is this related to forecasting?

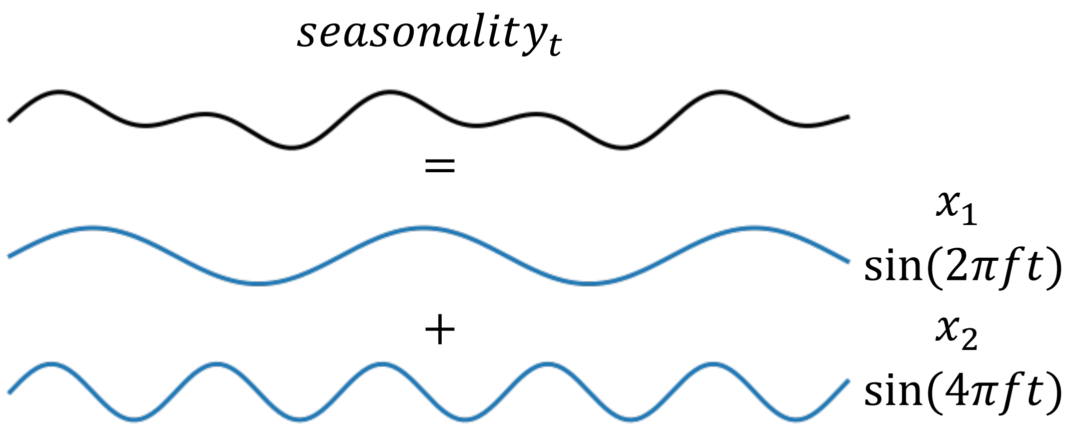
- We can use Fourier terms (sine & cosines) as features in a linear model to achieve this.

- Linear model:

$$
\begin{aligned}
y_t &= \beta_0 \ + \beta_1 x_1 + \beta_2 x_2 + \cdots \\
&= \beta_0 \ + \beta_1 \sin(2\pi f t) + \beta_2 \sin(4\pi f t) + \ \ldots
\end{aligned}
$$

- The Fourier coefficients, A_n and B_n, then become coefficients in the linear model, β_n, that are learned.

- The number of Fourier terms, N, is a hyperparameter controlling how complex the fitted seasonality is.
    - N is too large: overfit to noise.
    - N is too small: underfit the true seasonal pattern.



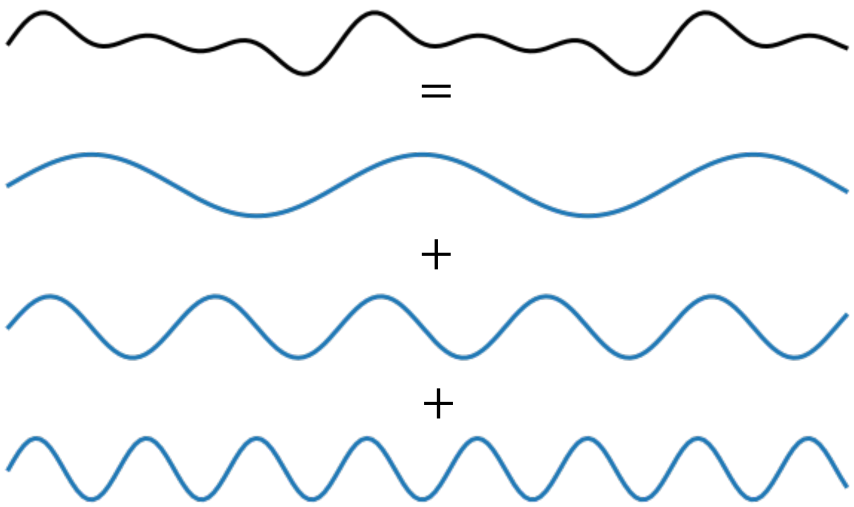
$N = 2$

How is this related to forecasting?

- We can use Fourier terms (sine & cosines) as features in a linear model to achieve this.
- Linear model:

$$\begin{aligned} y_t &= \beta_0 \ + \beta_1 x_1 + \beta_2 x_2 + \cdots \\ &= \beta_0 \ + \beta_1 \sin(2\pi f t) + \beta_2 \sin(4\pi f t) + \ \ldots \end{aligned}$$

- The Fourier coefficients, A_n and B_n, then become coefficients in the linear model, β_n, that are learned.
- The number of Fourier terms, N, is a hyperparameter controlling how complex the fitted seasonality is.
    - N is too large: overfit to noise.
    - N is too small: underfit the true seasonal pattern.



$N = 3$