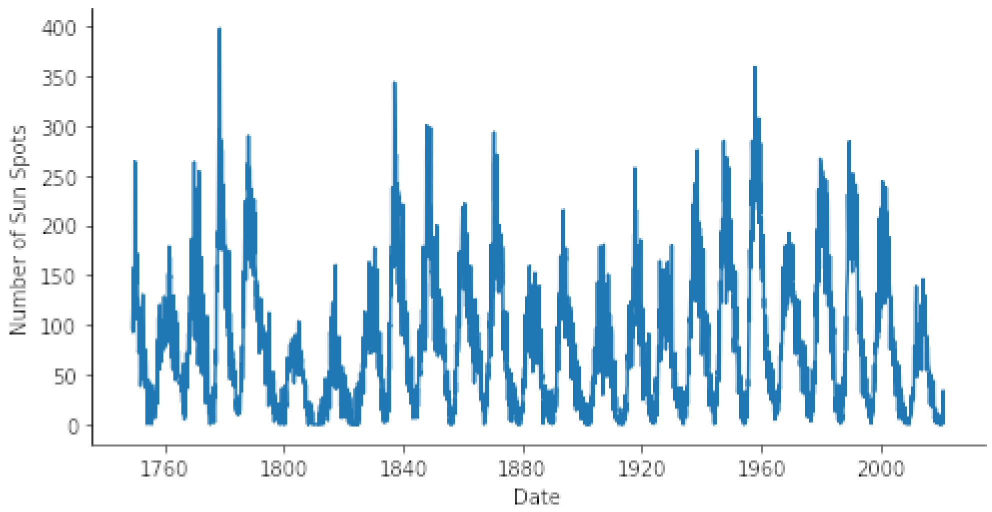
UNIVERSITY OF BIRMINGHAM

Visualisation

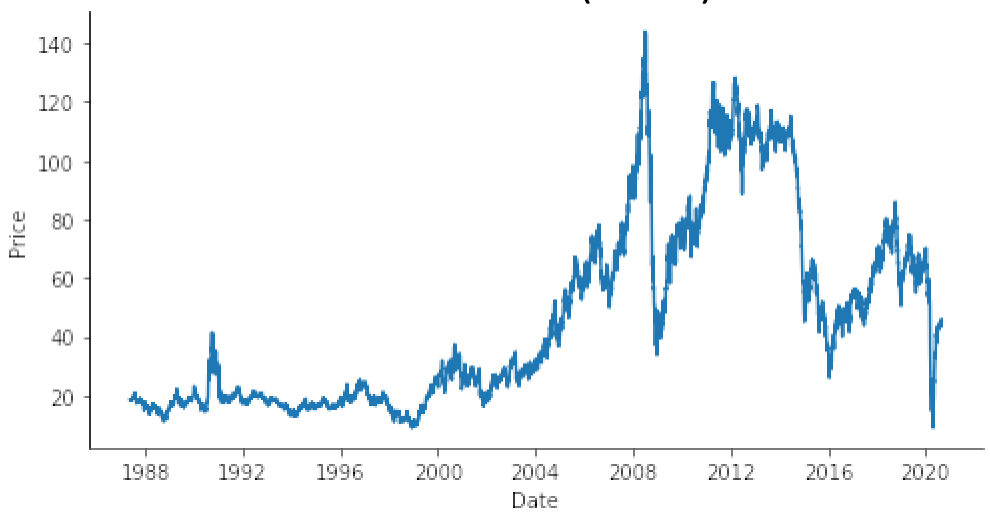
Week 3

Smoothing and Convolution

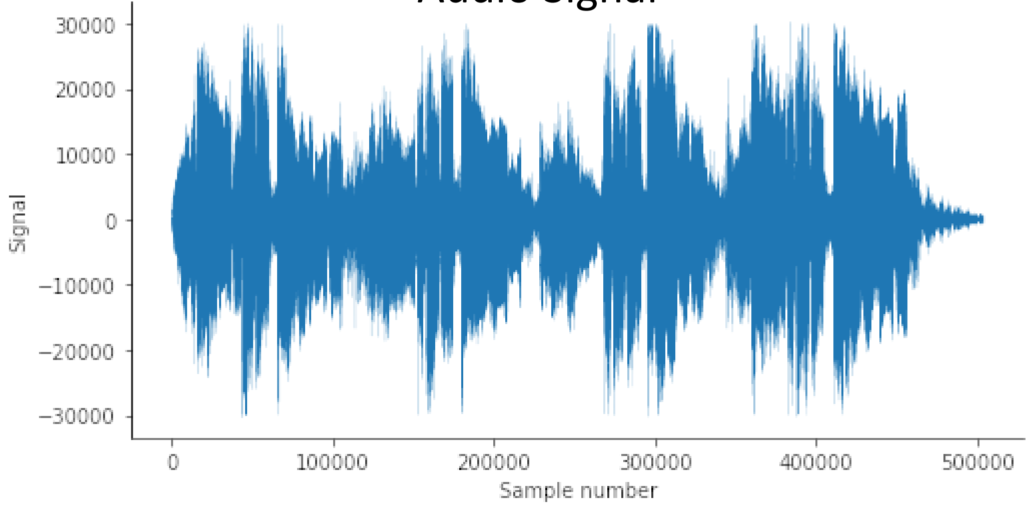
Number of Sunspots



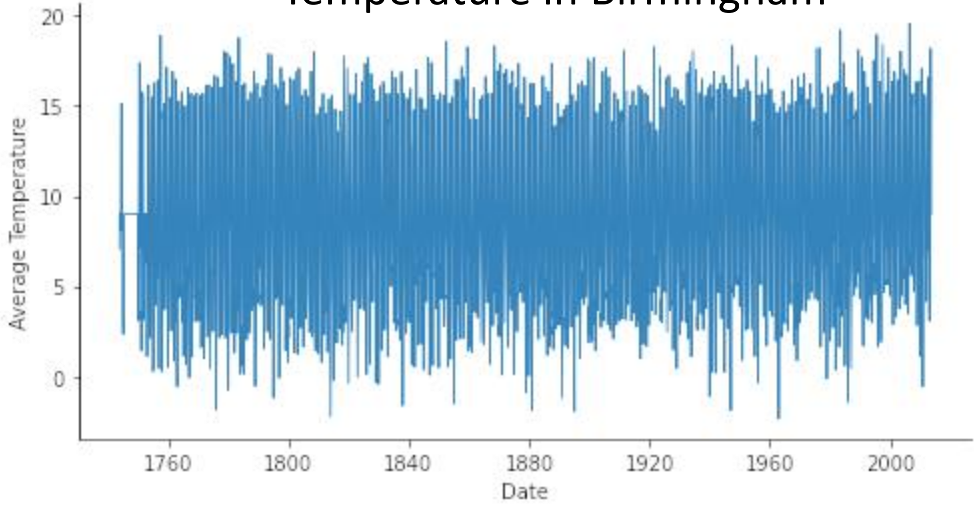
Oil Price (Brent)



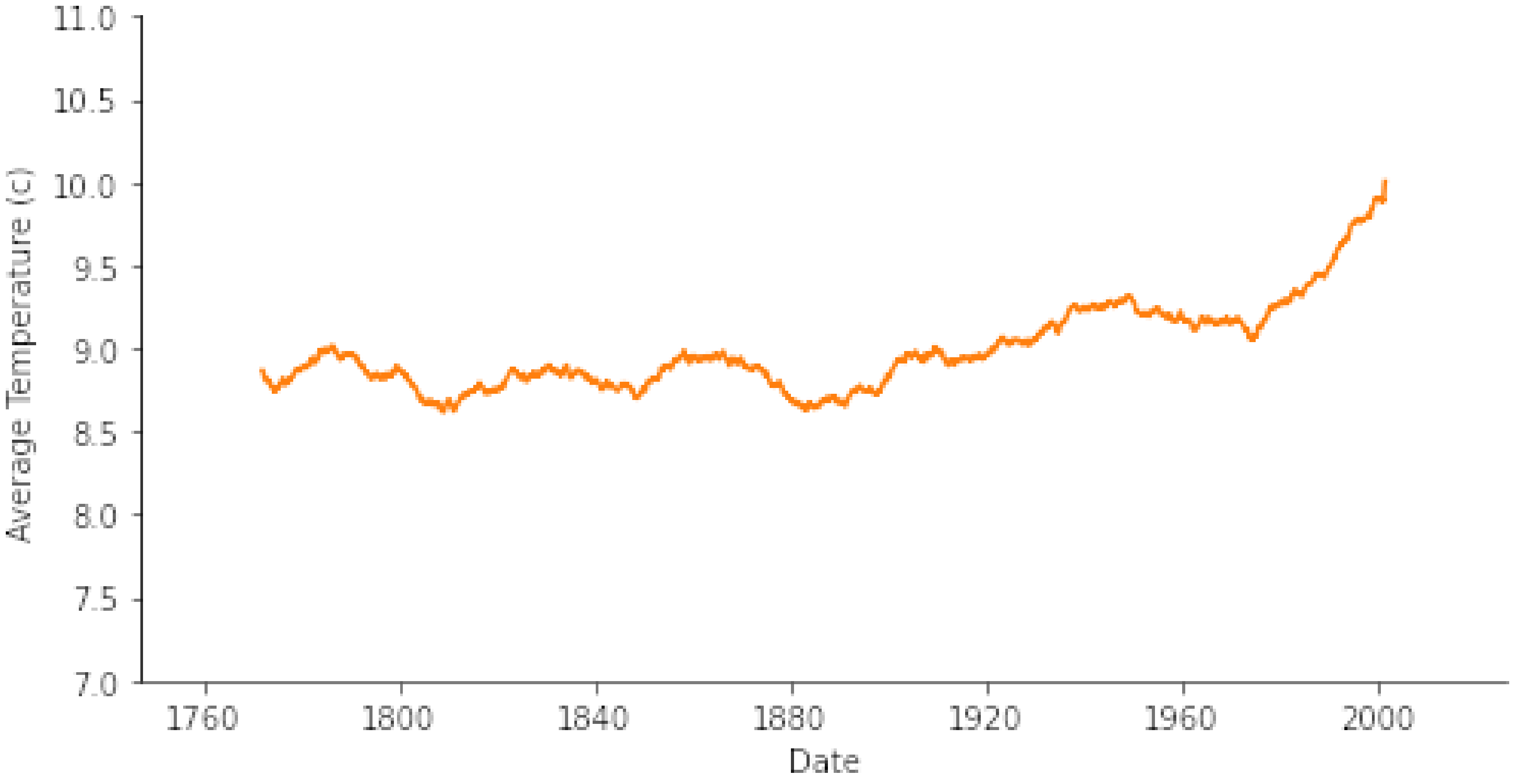
Audio Signal



Temperature in Birmingham



Smoothing

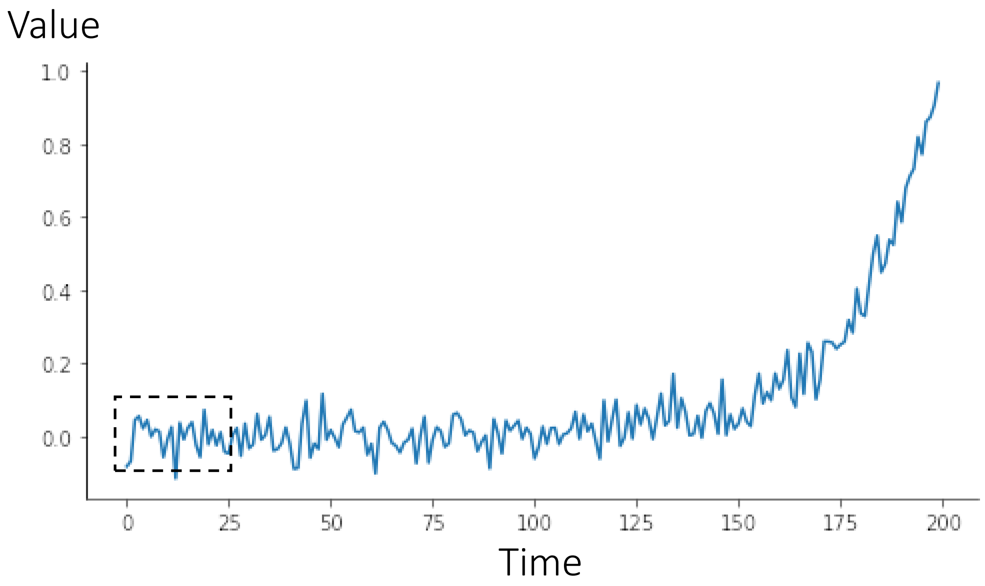


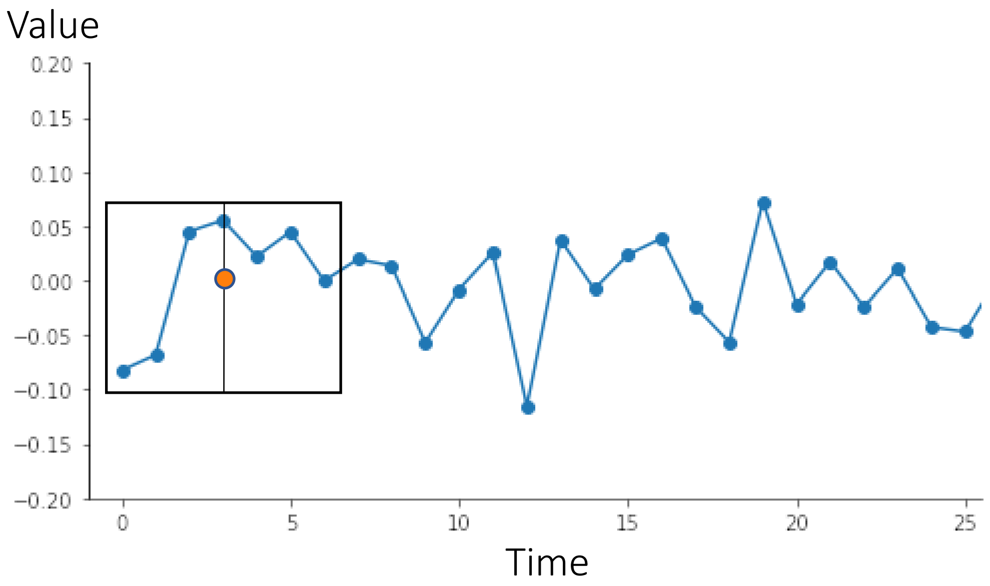
Temperature in Birmingham

Rolling Average Smoothing

aka Moving Average
aka Box(car) Smoothing

Rolling Average Smoothing



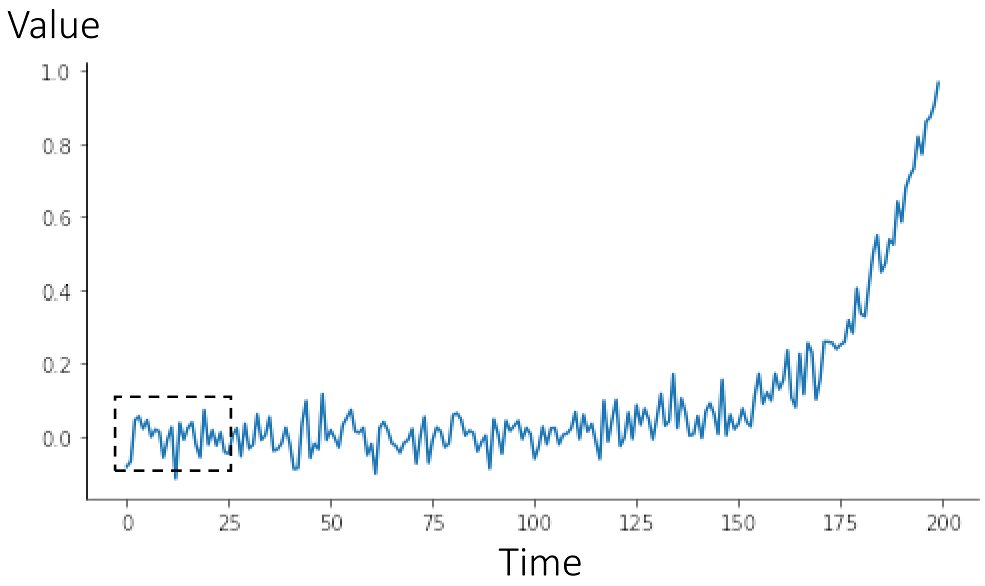


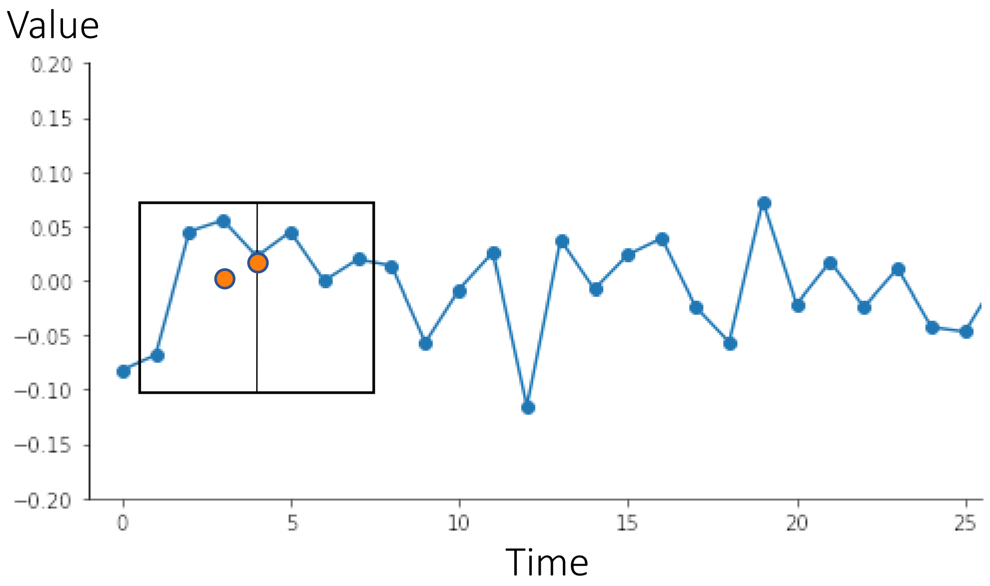
$$s_t = \frac{1}{w} \sum_{k=-m}^{m} x_{t+k}$$

Window size is w $= 2m + 1 = 7$

Here, $m = 3$

Rolling Average Smoothing



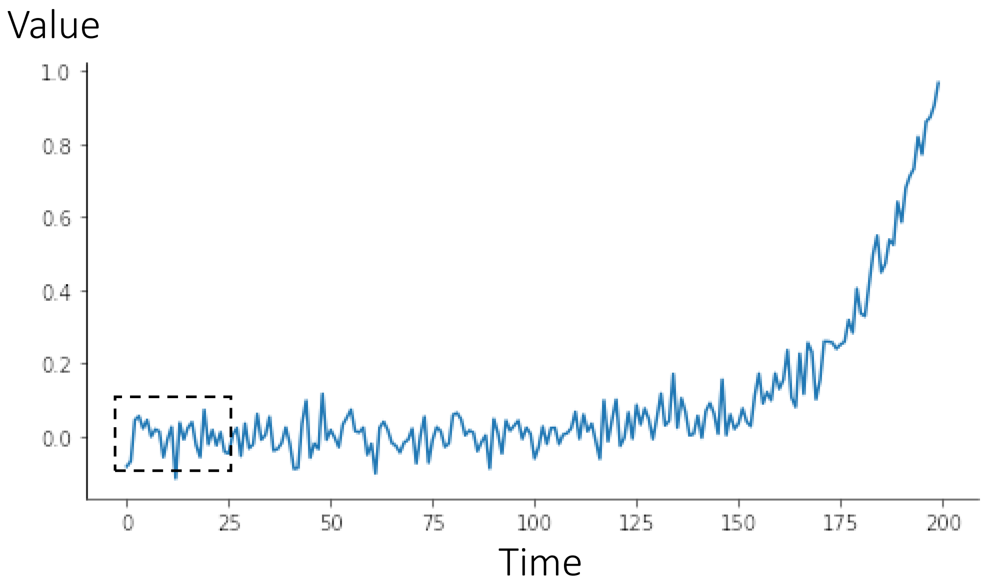


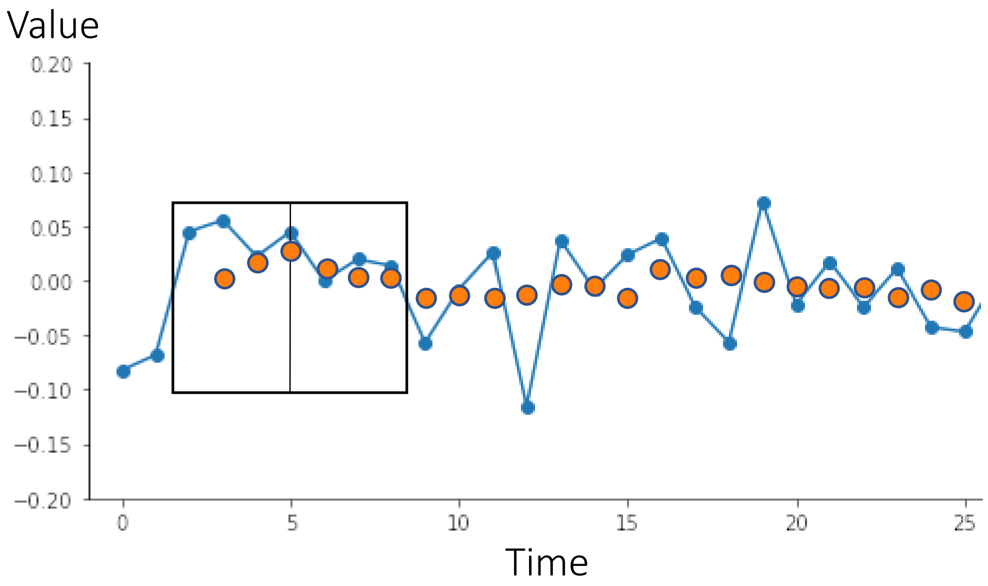
$$s_t = \frac{1}{w} \sum_{k=-m}^{m} x_{t+k}$$

Window size is w $= 2m + 1 = 7$

Here, $m = 3$

Rolling Average Smoothing



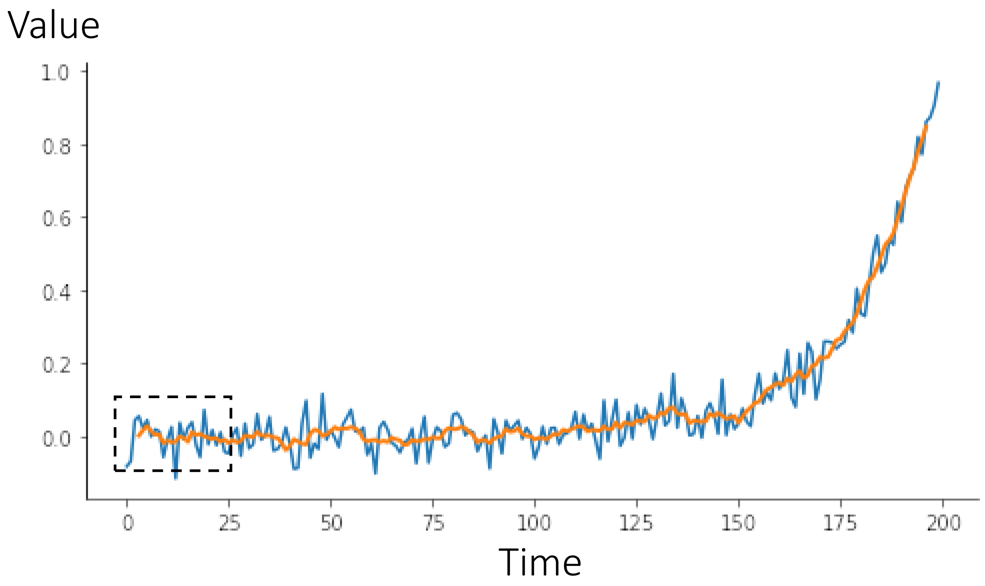


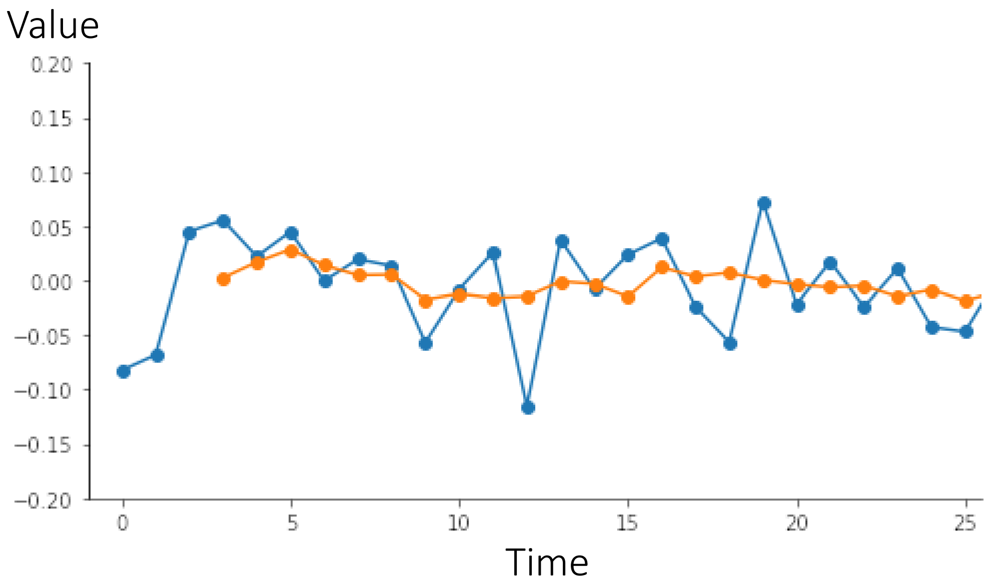
$$s_t = \frac{1}{w} \sum_{k=-m}^{m} x_{t+k}$$

Window size is w $= 2m + 1 = 7$

Here, $m = 3$

Rolling Average Smoothing



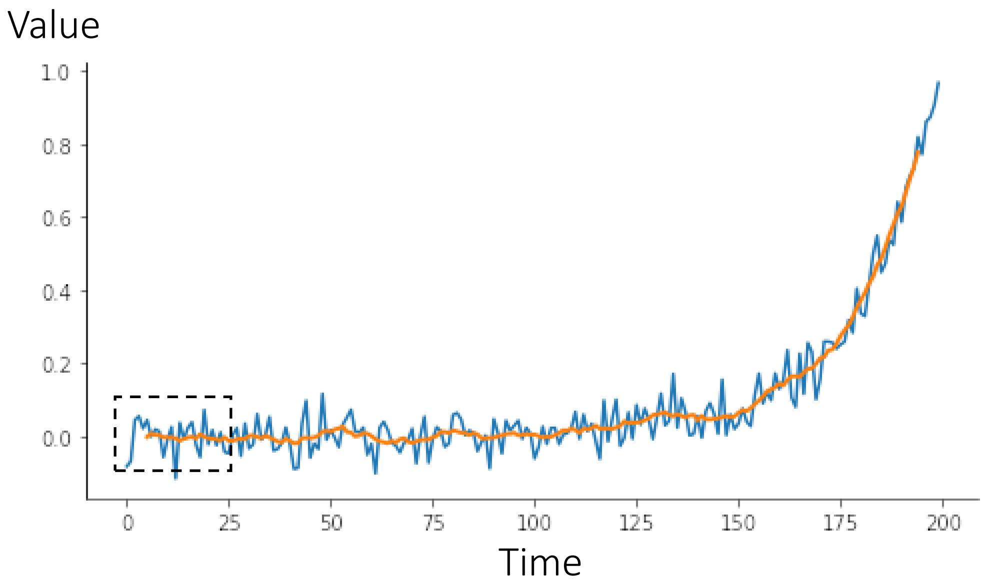


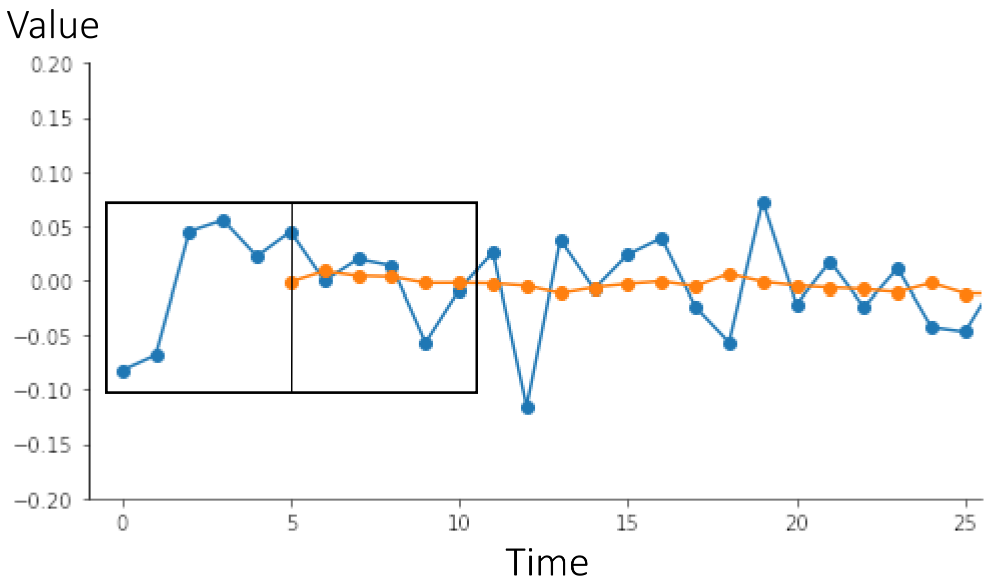
$$s_t = \frac{1}{w} \sum_{k=-m}^{m} x_{t+k}$$

Window size is w $= 2m + 1$

Here, $m = 3$

Rolling Average Smoothing



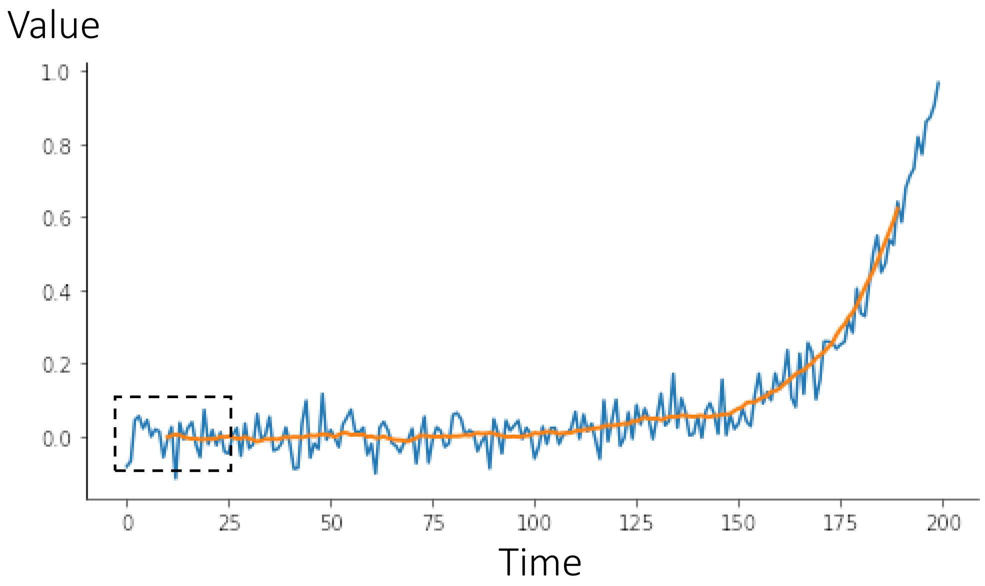


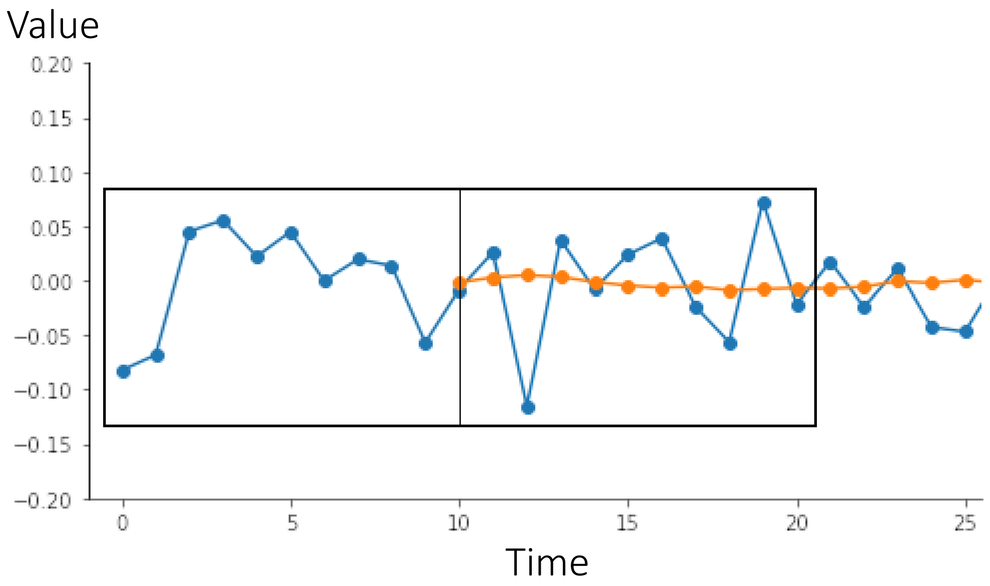
$$s_t = \frac{1}{w} \sum_{k=-m}^{m} x_{t+k}$$

Window size is $\mathrm{w} = 2m + 1 = 11$

Here, $m = 5$

Rolling Average Smoothing





$$s_t = \frac{1}{w} \sum_{k=-m}^{m} x_{t+k}$$

Larger window gives:

- Smoother result
- Larger areas without estimate

Window size is w $= 2m + 1 = 21$

Here, $m = 10$

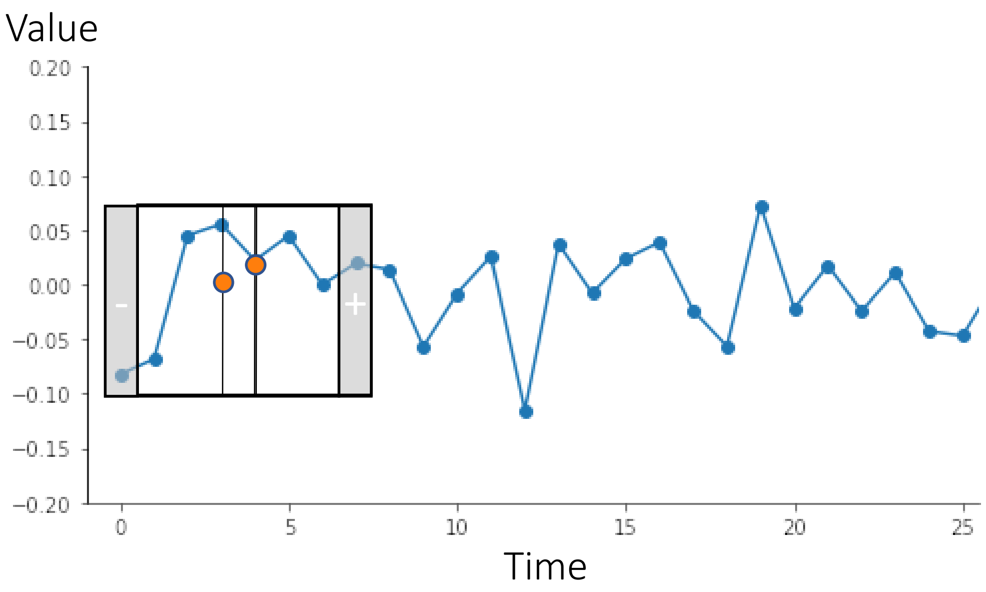
Rolling Average Smoothing - Computation

Naïve implementation:

$$s_t = \frac{1}{w}\sum_{k=-m}^{m} x_{t+k}$$

$$= \sum_{k=-n}^{n} \frac{x_{t+k}}{w}$$

$O(wn)$



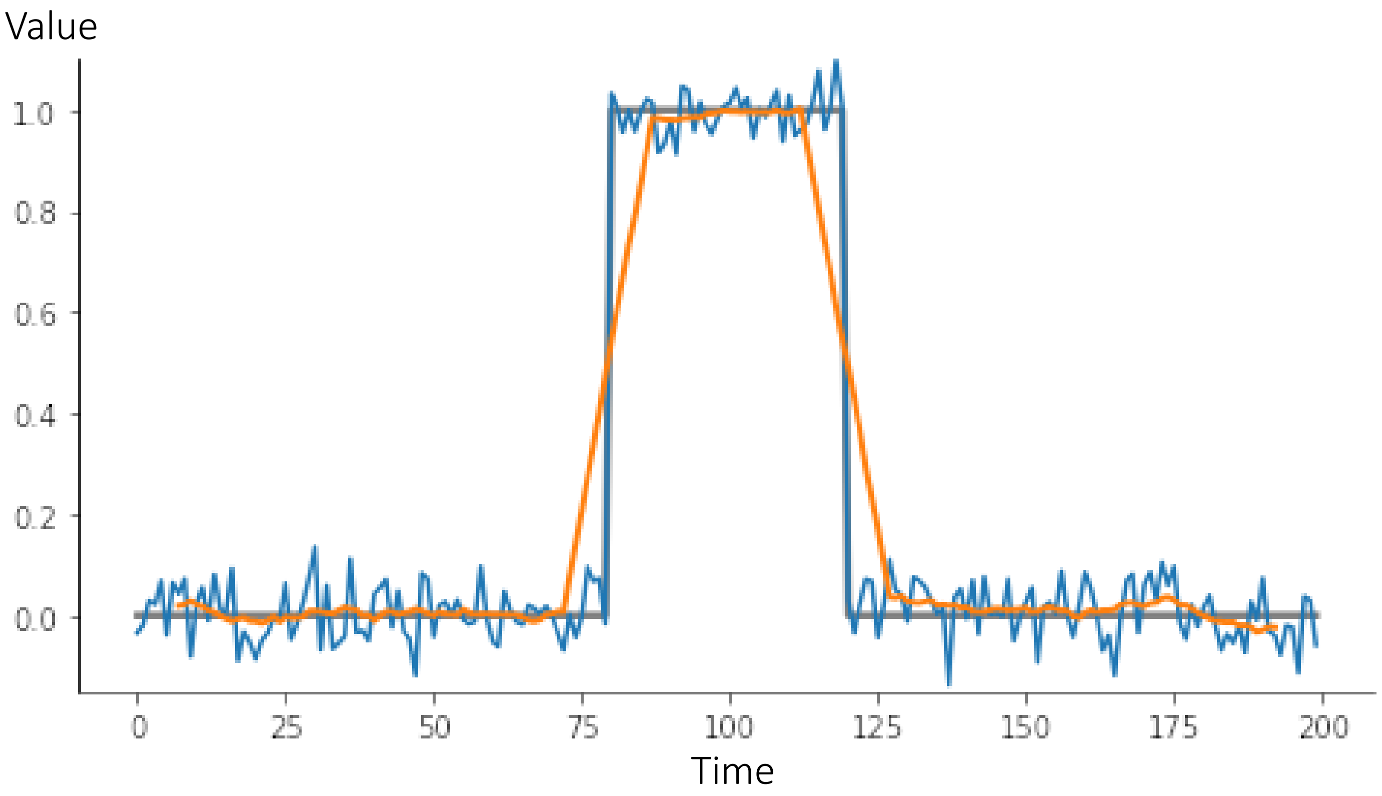
Recursive implementation:

$$s_t = s_{t-1} - \frac{x_{t-k}}{w} + \frac{x_{t+k}}{w}$$

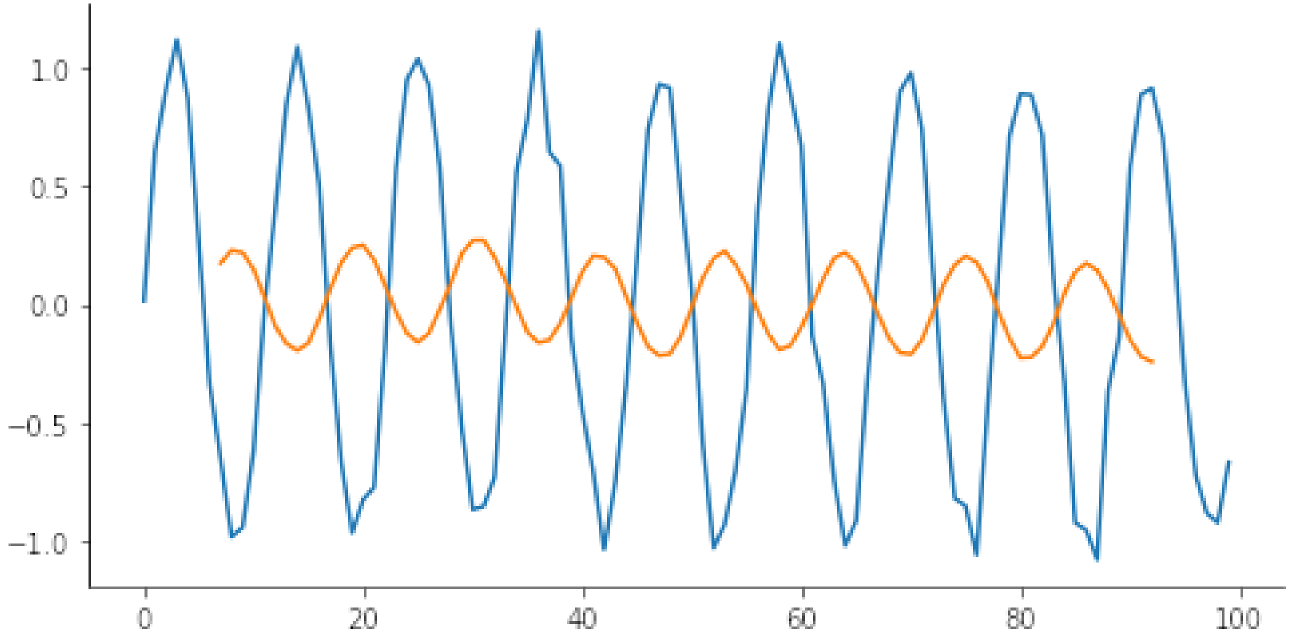
$O(n)$

Very efficient!

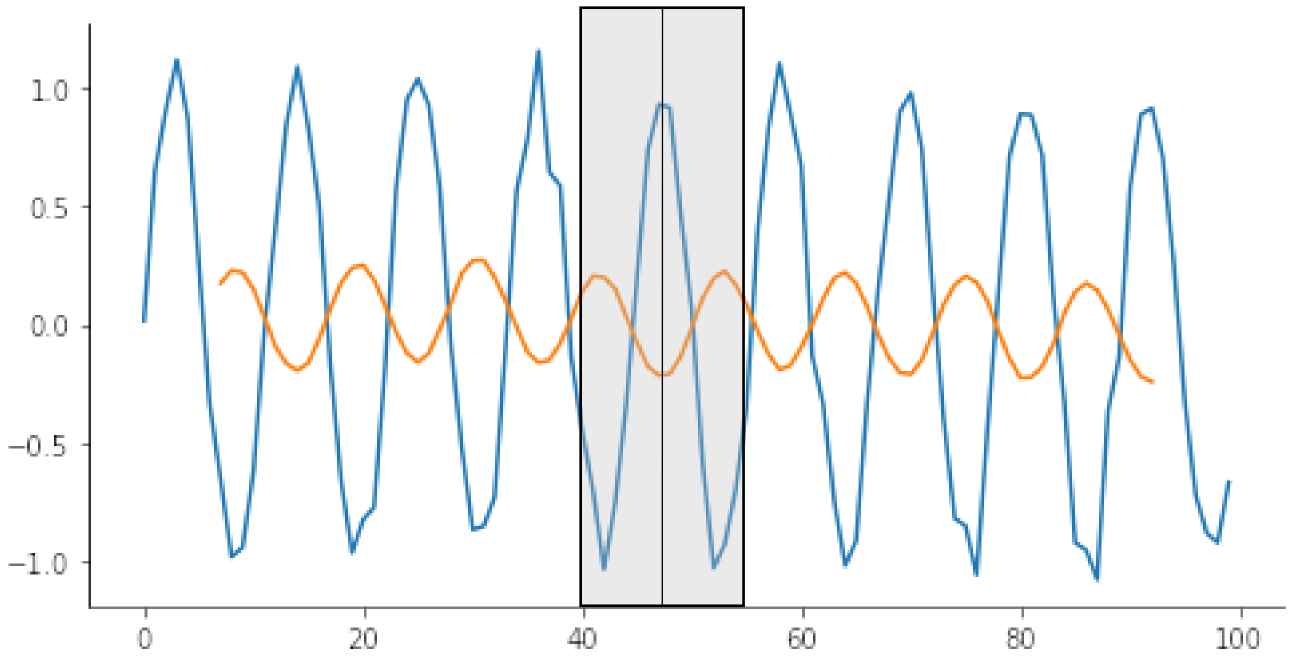
Rolling Average Smoothing - Downsides



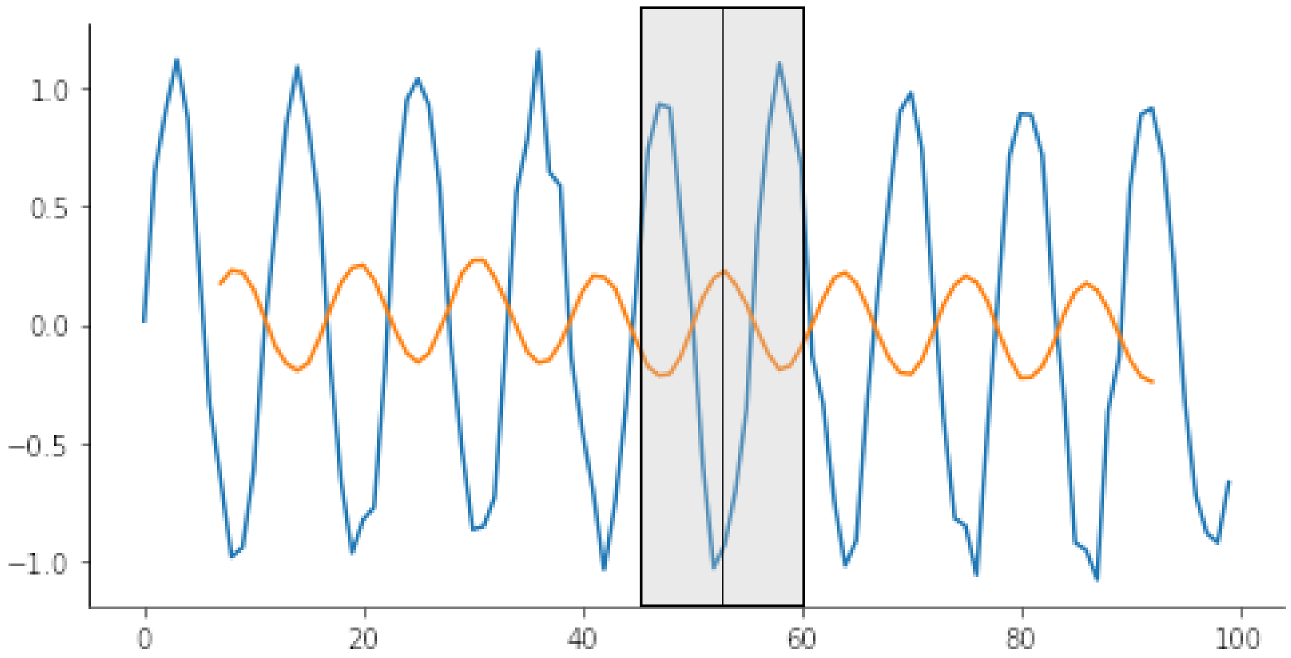
Rolling Average Smoothing - Downsides



Strange behaviour depending on frequency and window size.

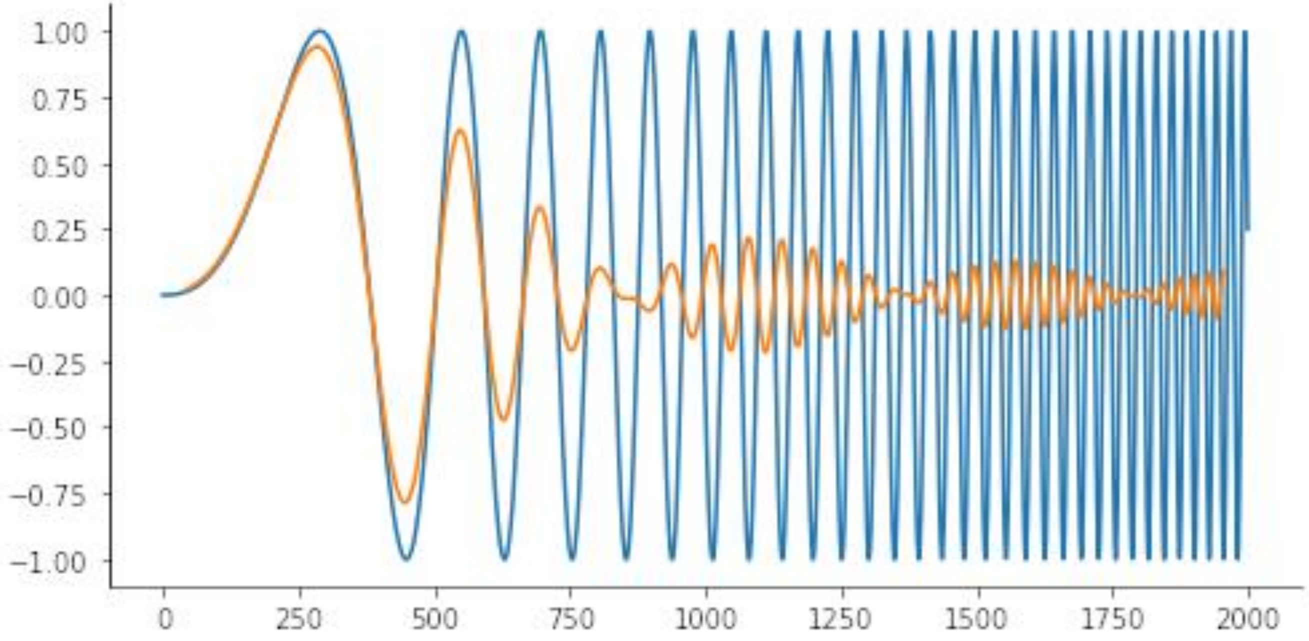


Strange behaviour depending on frequency and window size

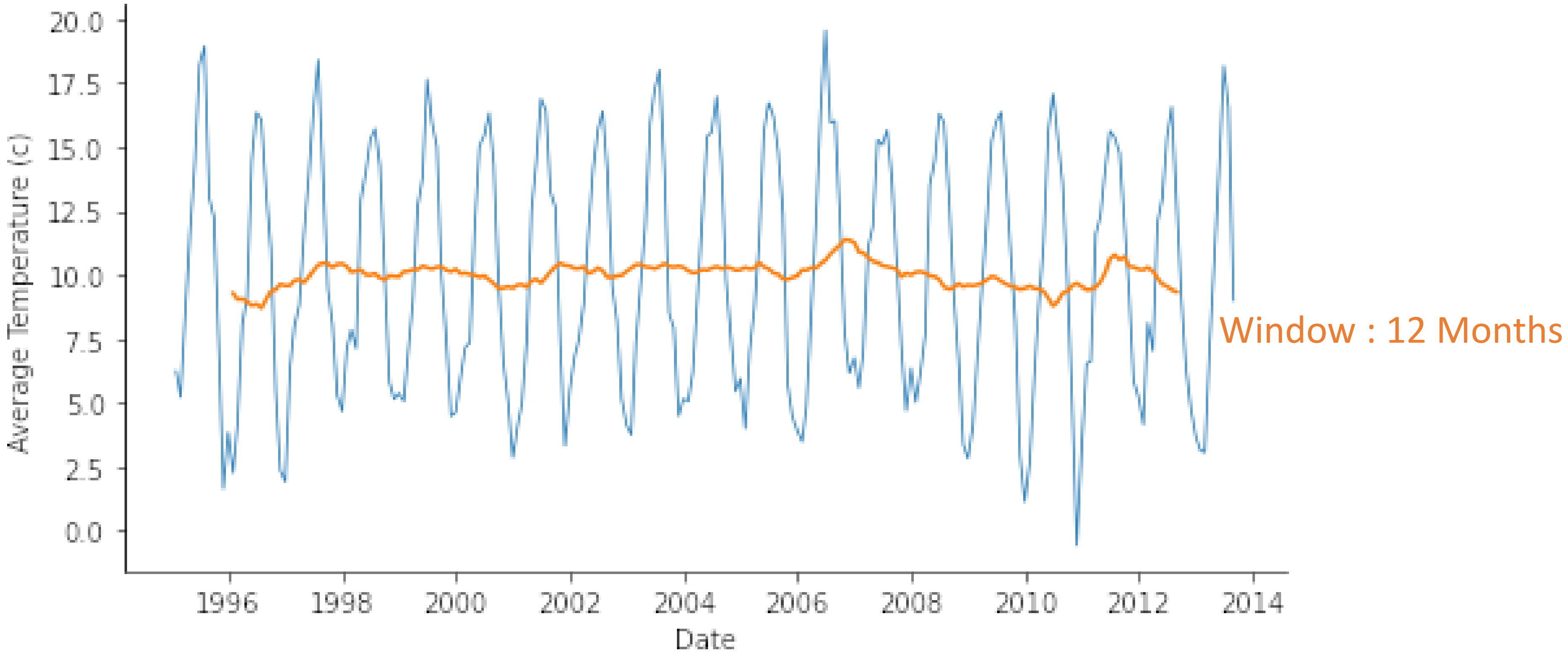


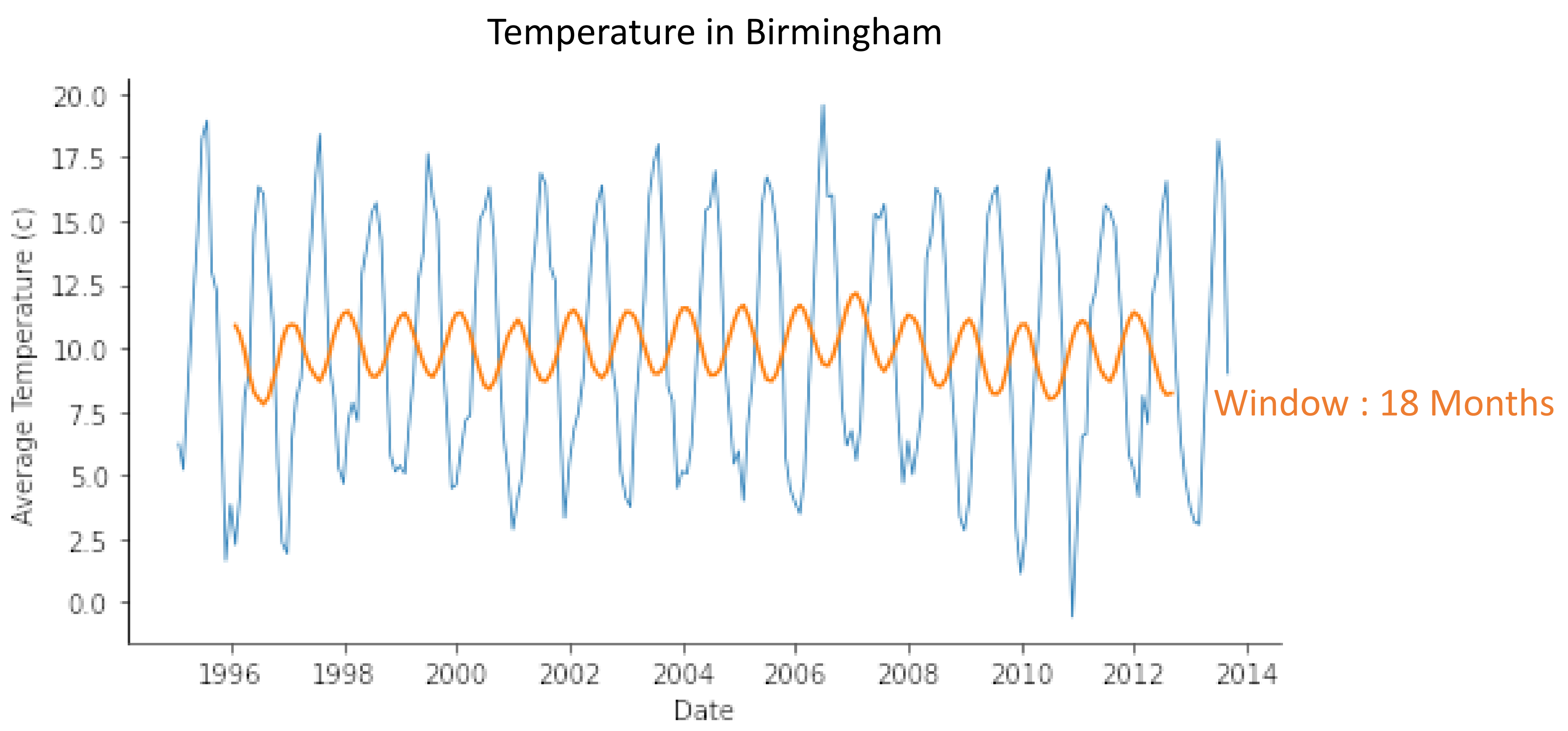
Strange behaviour depending on frequency and window size

Rolling Average Smoothing - Downsides



Temperature in Birmingham



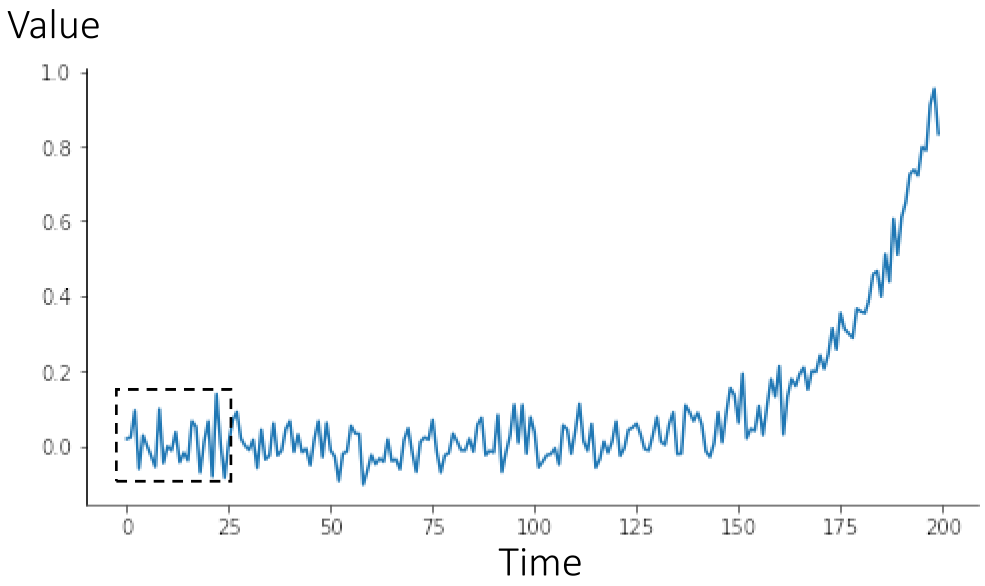


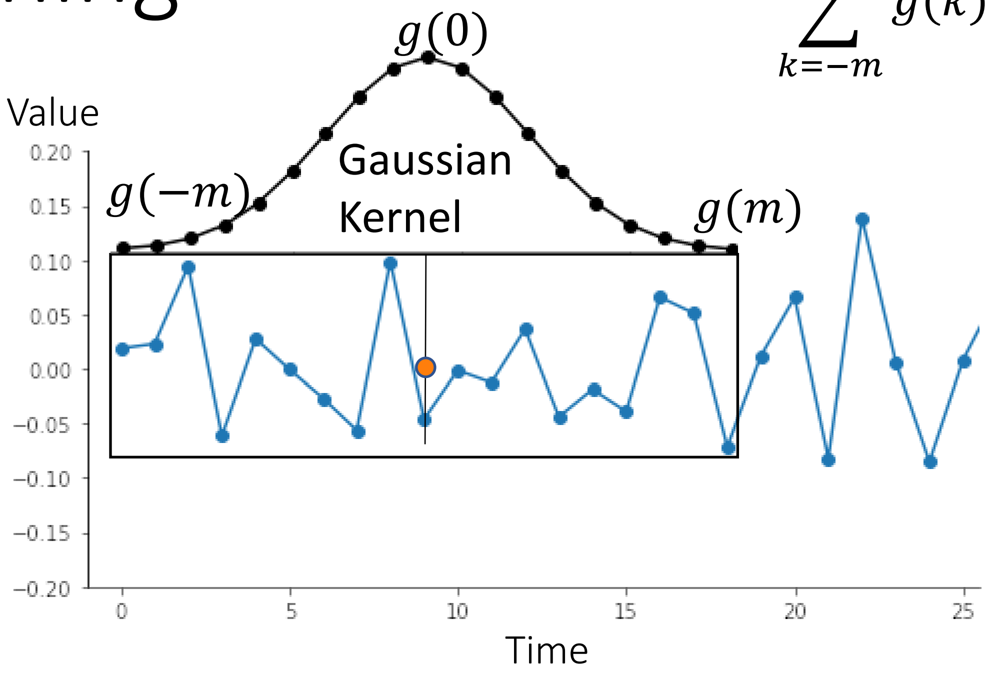
Temperature in Birmingham
Average Temperature (c)
20.0
17.5
15.0
12.5
10.0
7.5
5.0
2.5
0.0
1996
1998
2000
2002
2004
2006
2008
2010
2012
2014
Date
Window : 18 Months

Weighted Average Smoothing

Weighted Average Smoothing

$$\sum_{k=-m}^{m} g(k) = 1$$

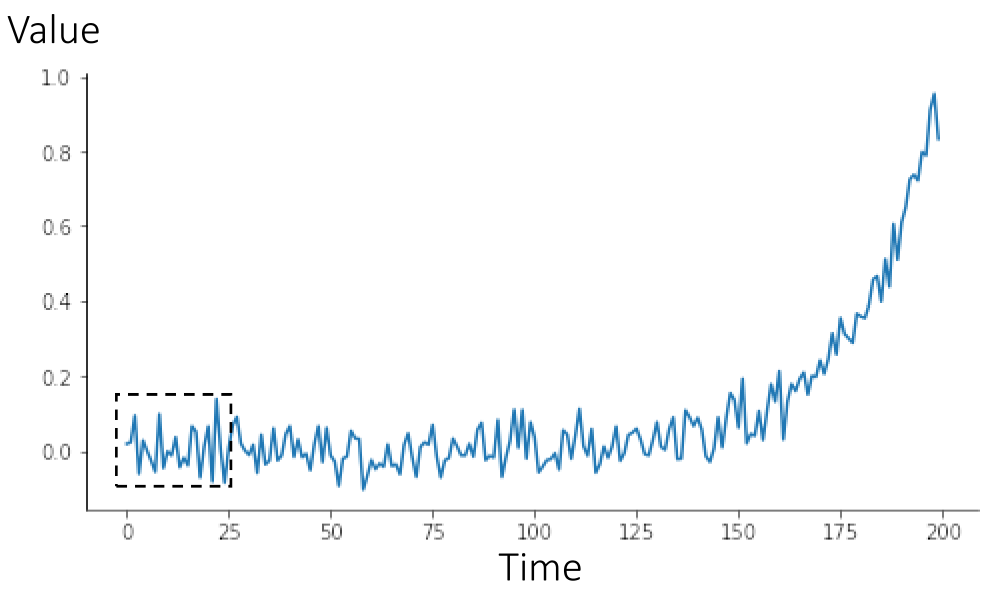


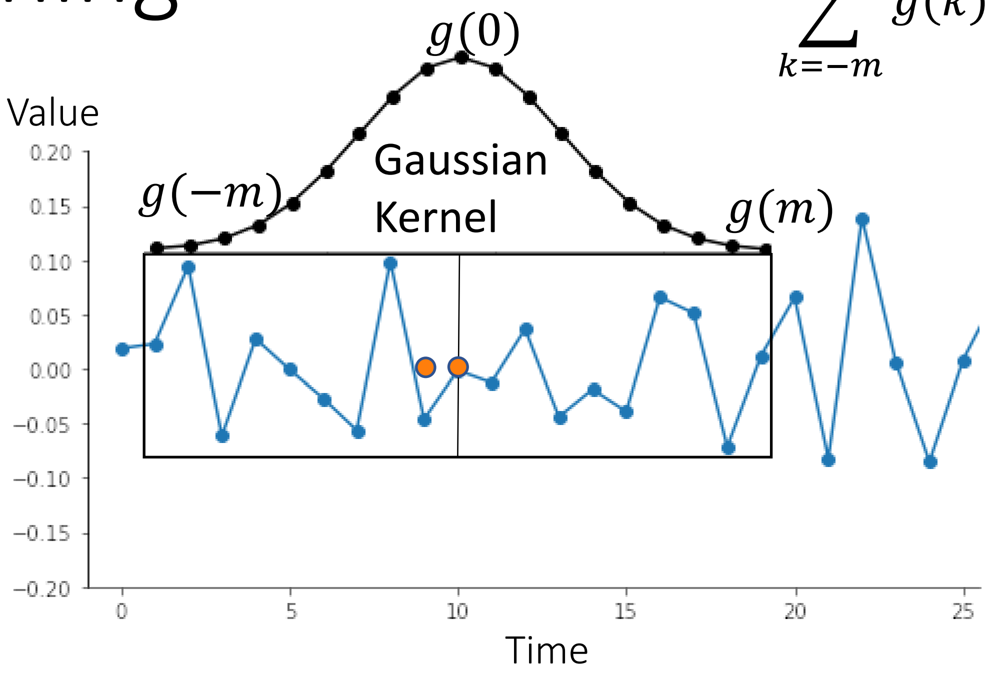


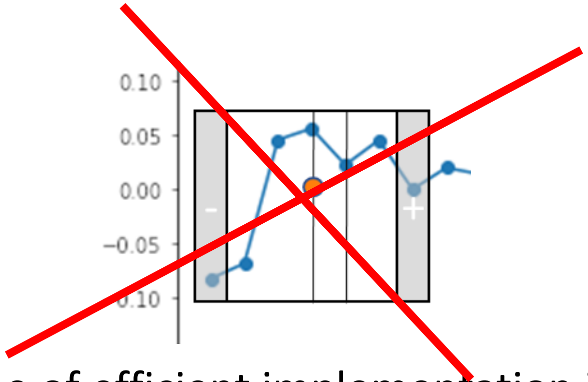
$$s_t = \sum_{k=-m}^{m} x_{t+k}\, g(k)$$

Weighted Average Smoothing

$$\sum_{k=-m}^{m} g(k) = 1$$



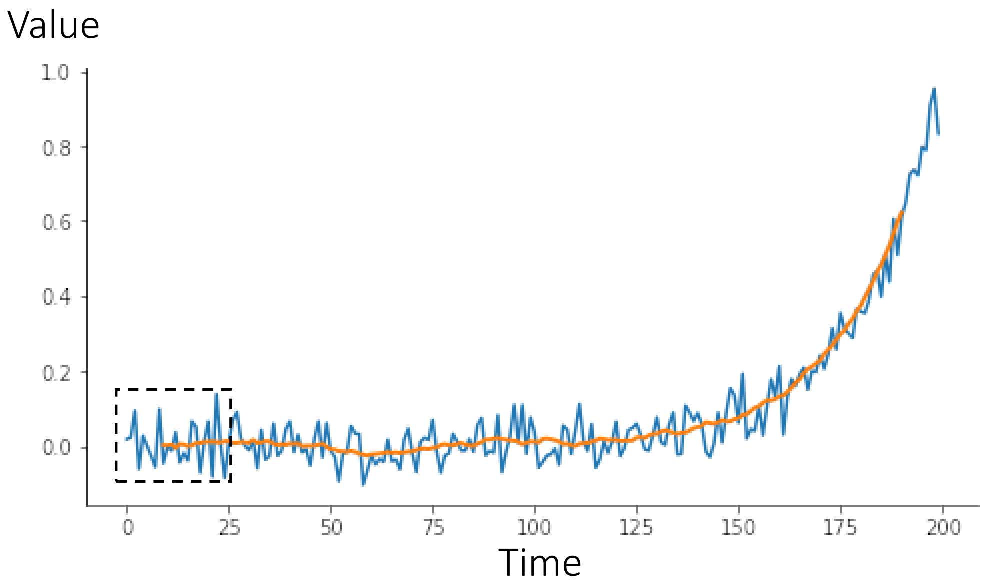


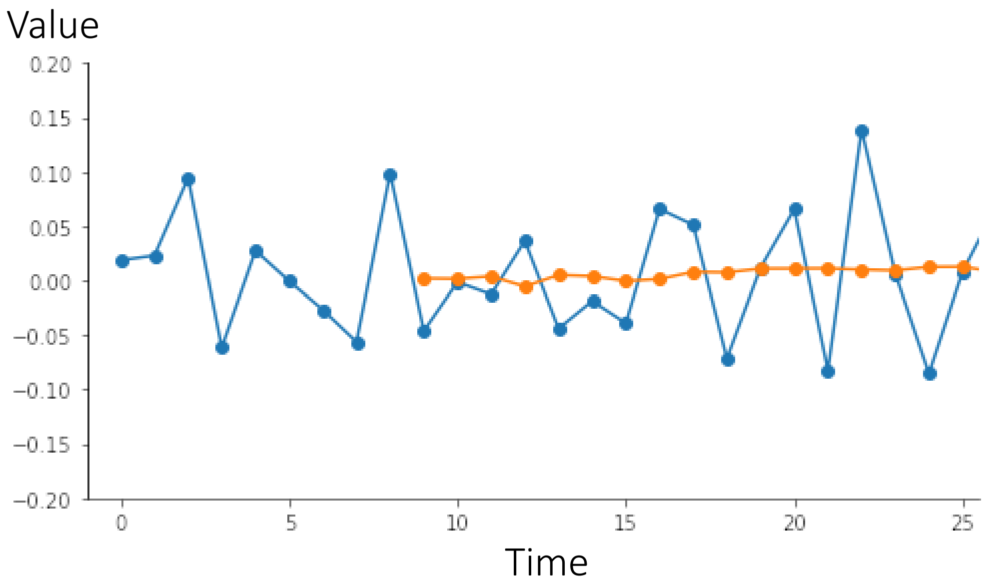


This type of efficient implementation is not possible.

$$s_t = \sum_{k=-m}^{m} x_{t+k}\, g(k)$$

Weighted Average Smoothing

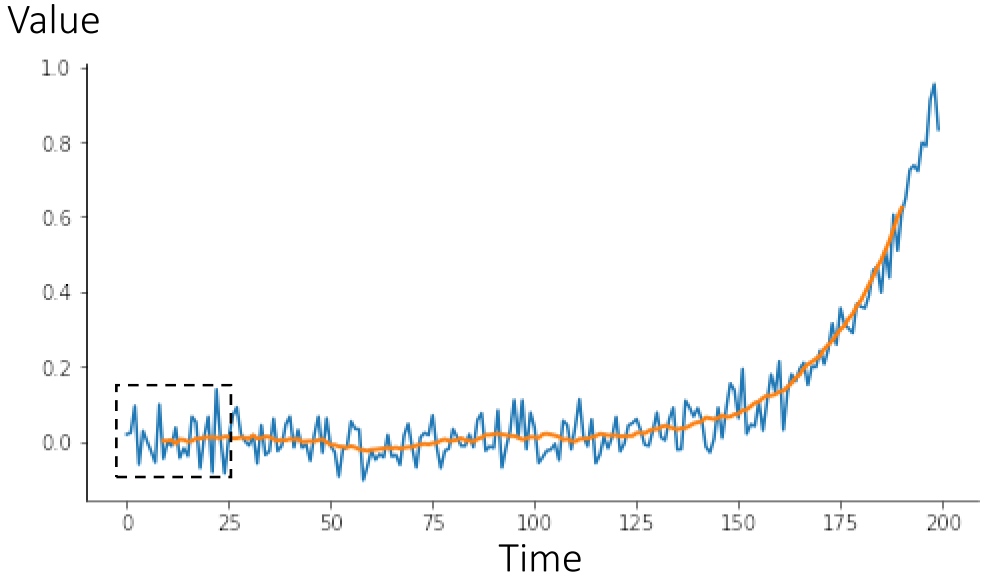


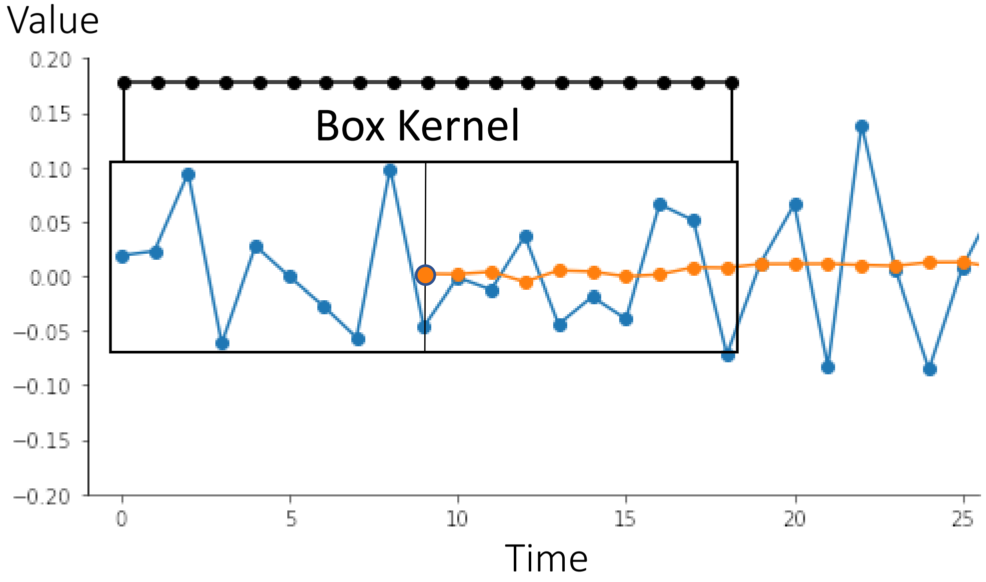


$$s_t = \sum_{k=-m}^{m} x_{t+k}\, g(k)$$

Weighted Average Smoothing

$$\sum_{k=-m}^{m} g(k) = 1$$

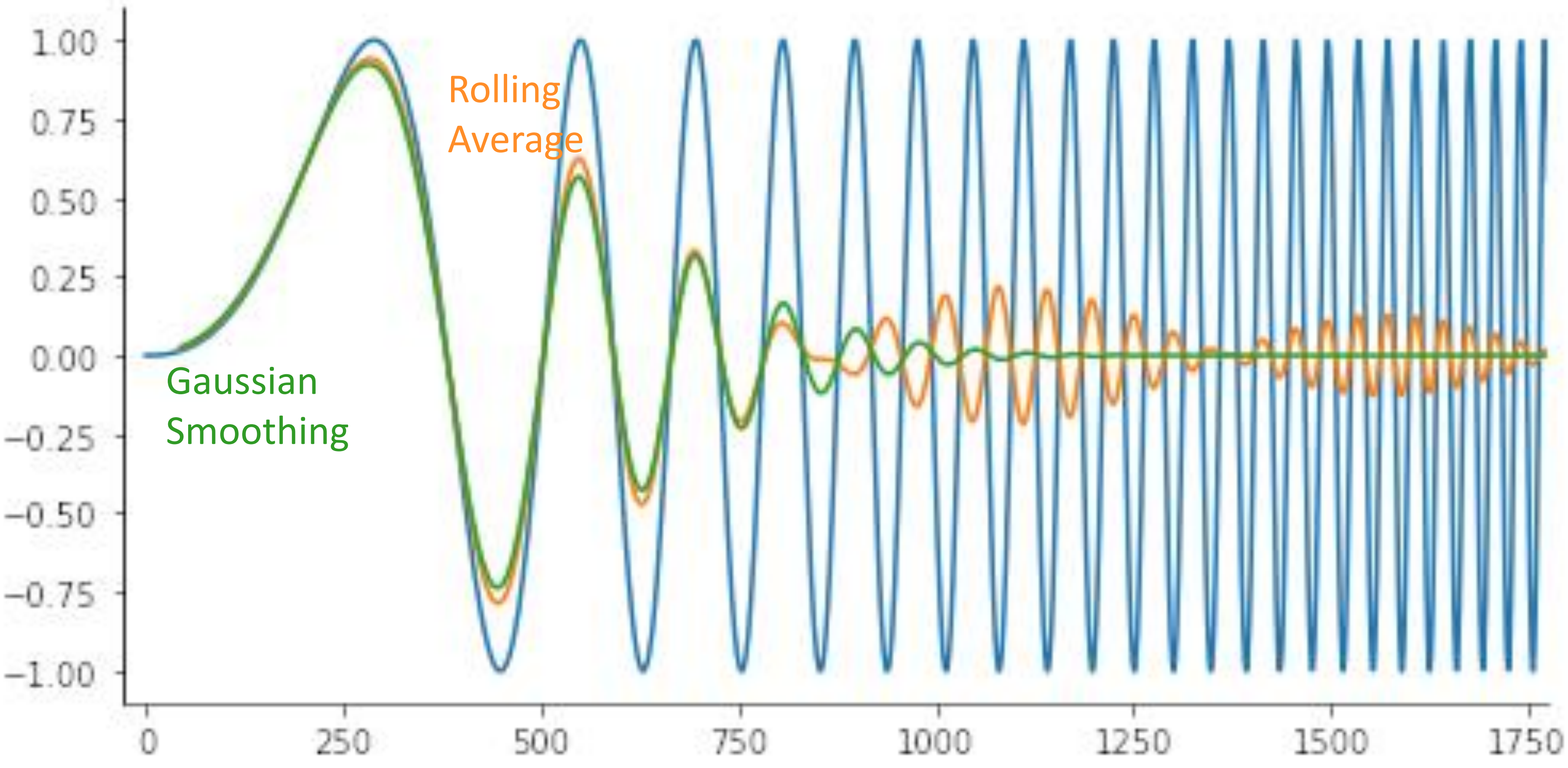




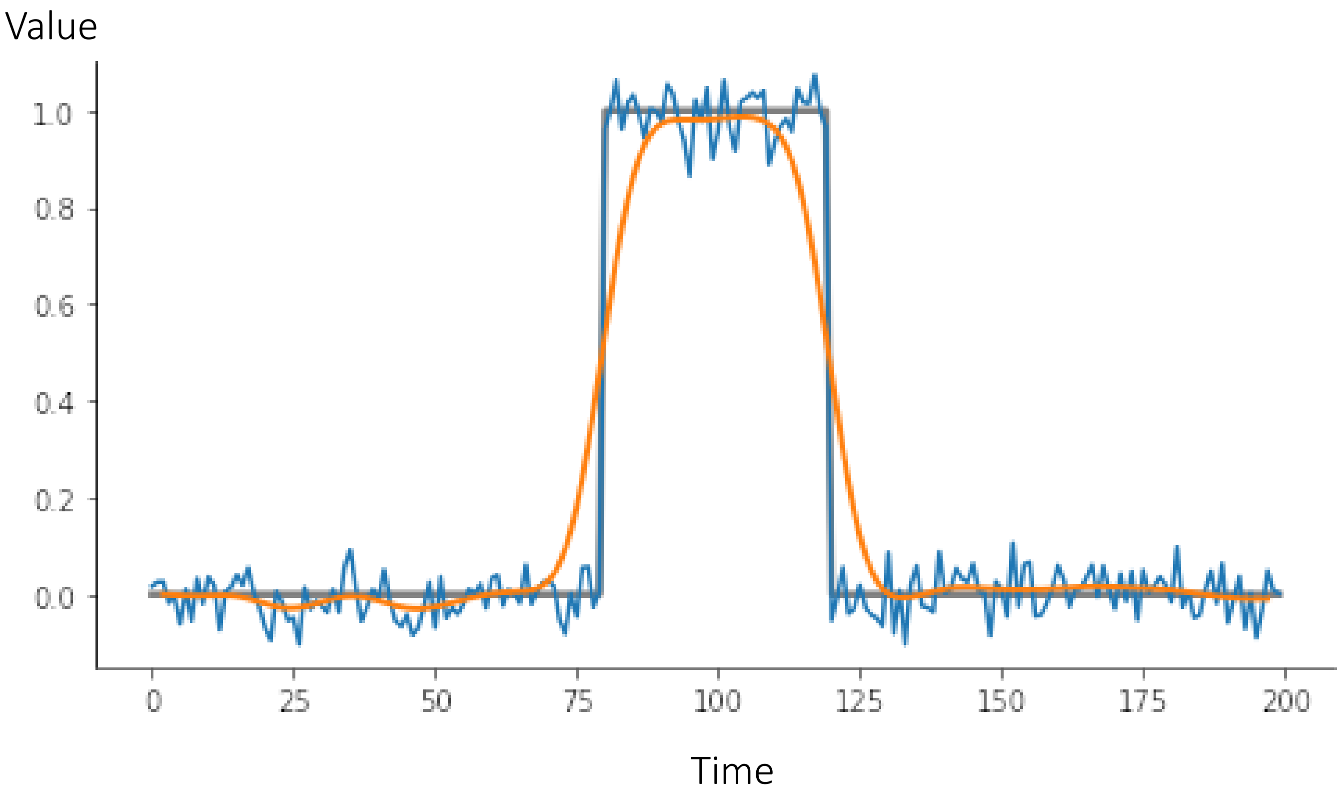
$$s_t = \sum_{k=-m}^{m} x_{t+k}\, g(k) \qquad g(k) = \frac{1}{w}$$

Rolling Average is a special case of Weighted Averaging.

Advantages of Gaussian Smoothing



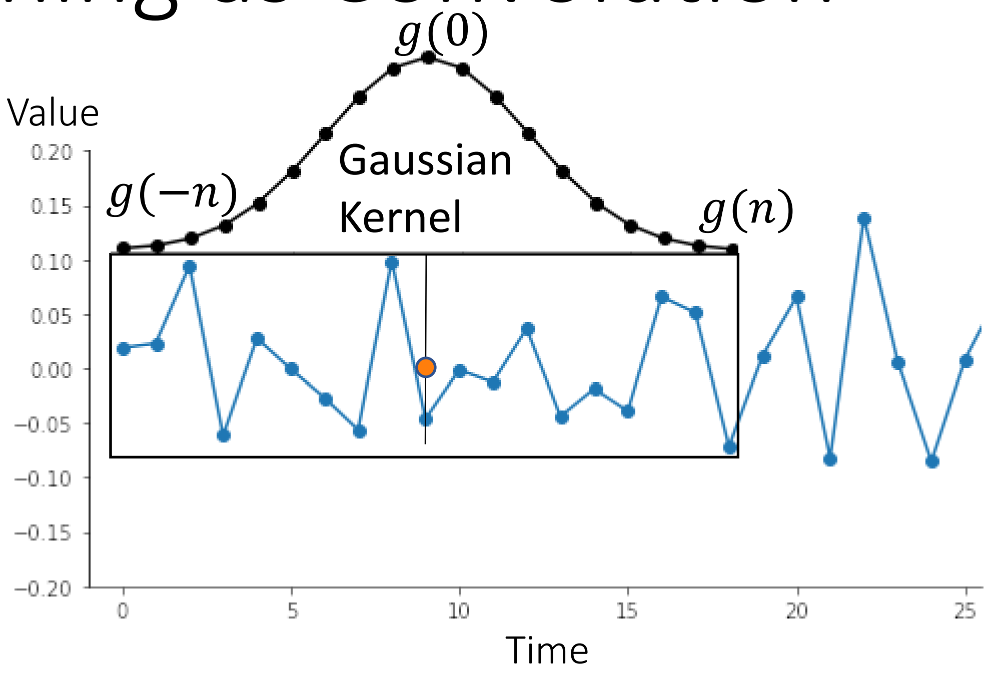
Gaussian Filter Downsides



Weighted Average Smoothing as Convolution

Definition of Convolution

$$(f * g)(t) = \int_{-\infty}^{\infty} f(t-k)g(k)dk$$



Discrete Version

$g(k) = 0$ outside $[-m, m]$

$$(f * g)(t) = \sum_{k=-\infty}^{\infty} f(t-k)\, g(k)$$

Flipped $g(k)$

$$s_t = \sum_{k=-m}^{m} x_{t+k}\, g(k)$$

Weighted averaging is a convolution

Convolutions - Properties

Commutativity

$$f * g = g * f$$

Associativity

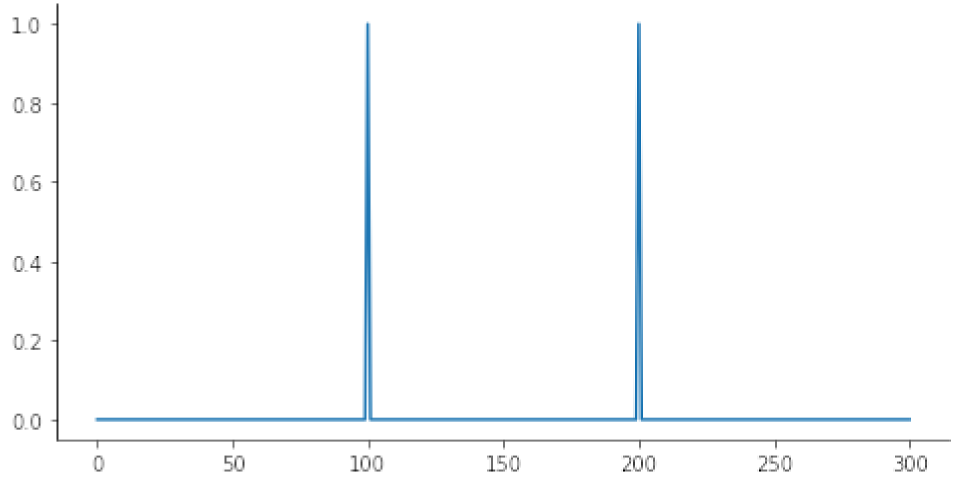
$$(f * g) * h = f * (g * h)$$

$$a(f * g) = (af) * g$$

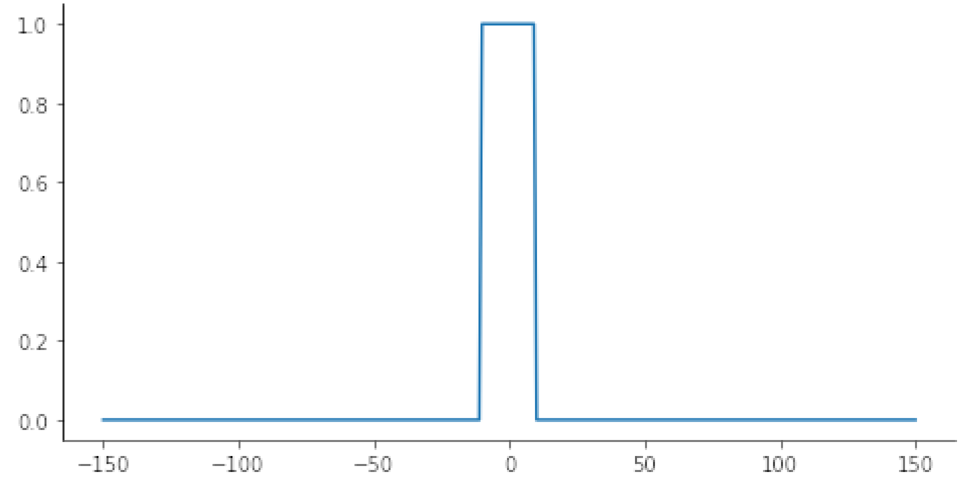
Distributivity

$$(f + g) * h = (f * h) + (g * h)$$

Convolutions - Examples

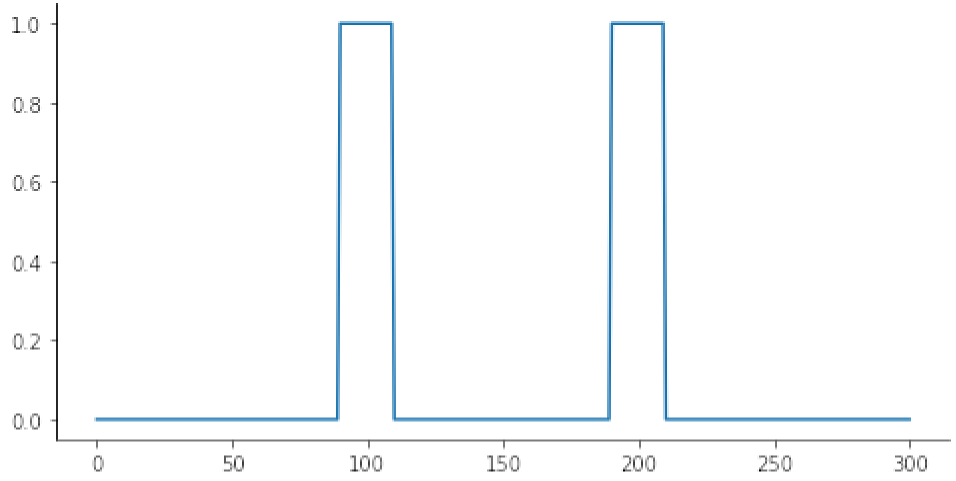


*

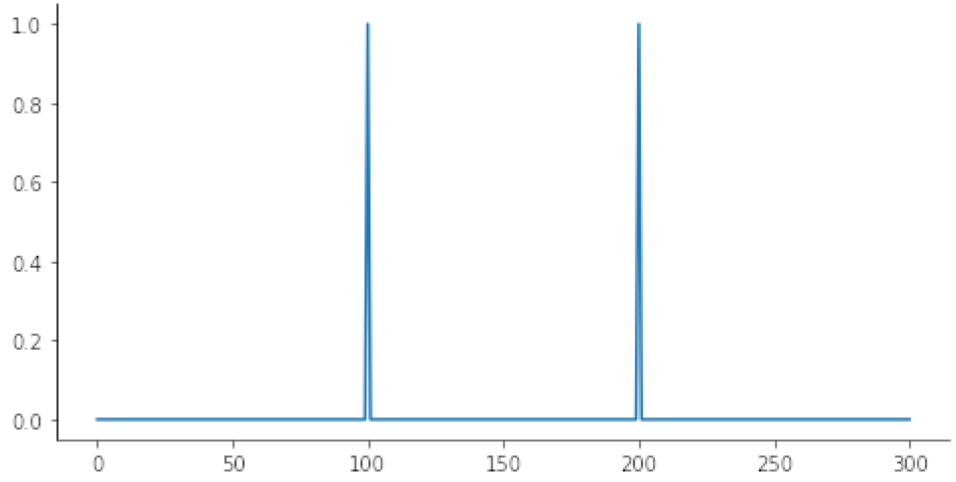


=

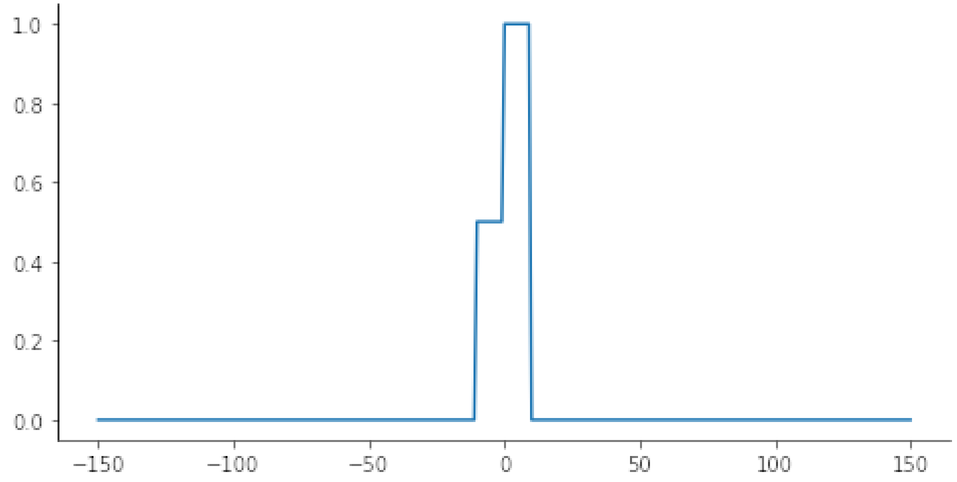
$$(f * g)(t) = \sum_{k=-\infty}^{\infty} f(t-k)\, g(k)$$



Convolutions - Examples

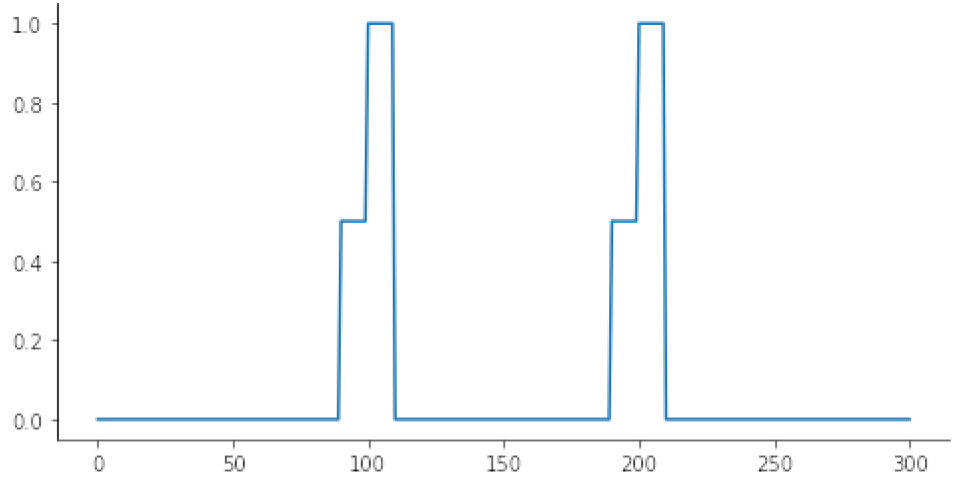


*

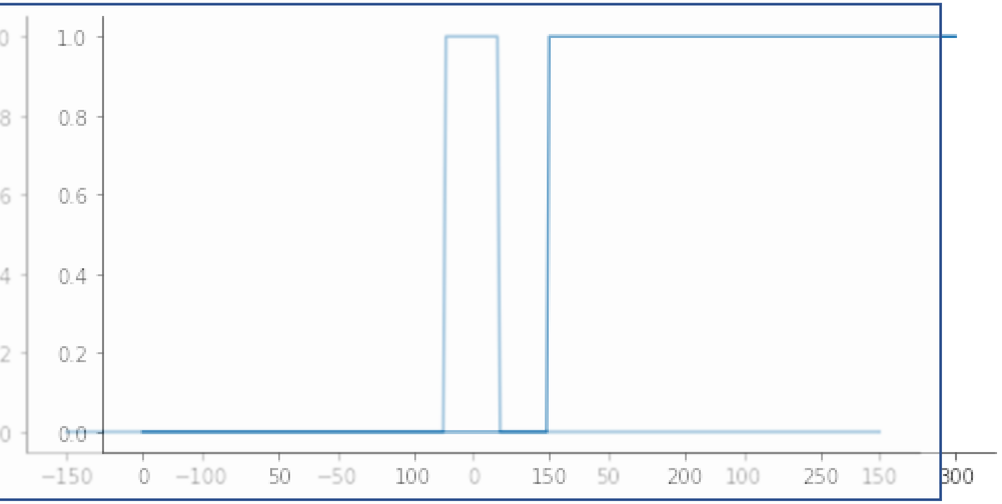


=

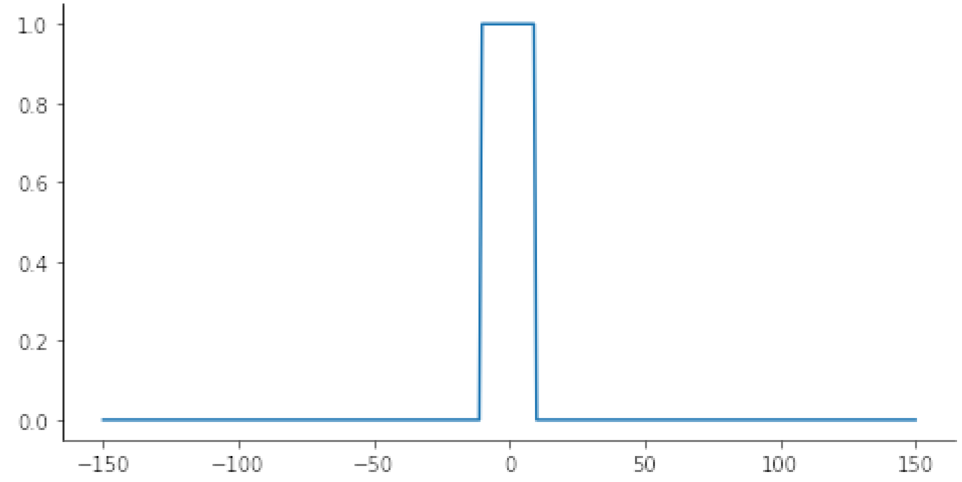
$$(f * g)(t) = \sum_{k=-\infty}^{\infty} f(t-k)\, g(k)$$



Convolutions - Examples

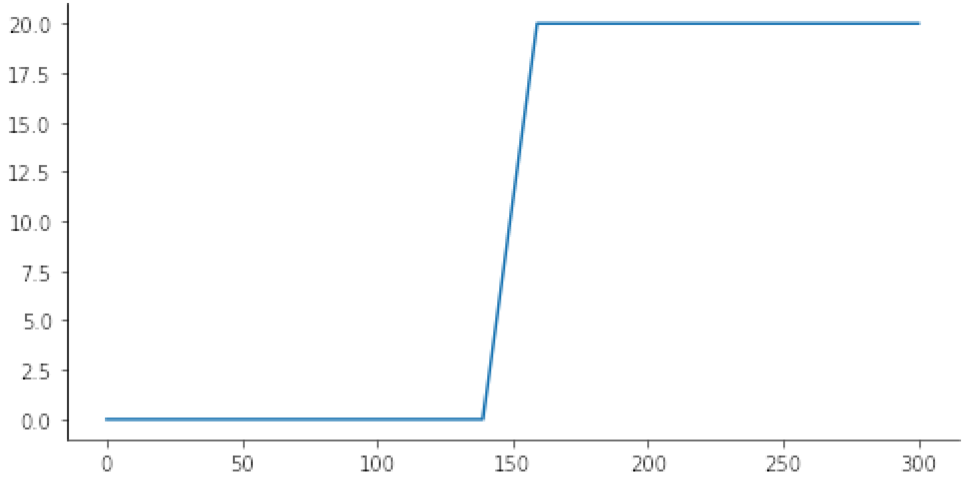


$*$

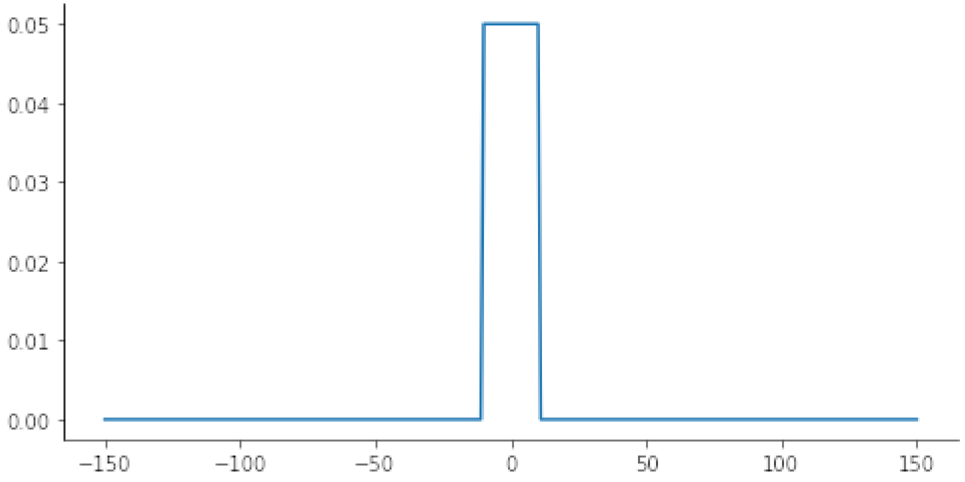


$=$

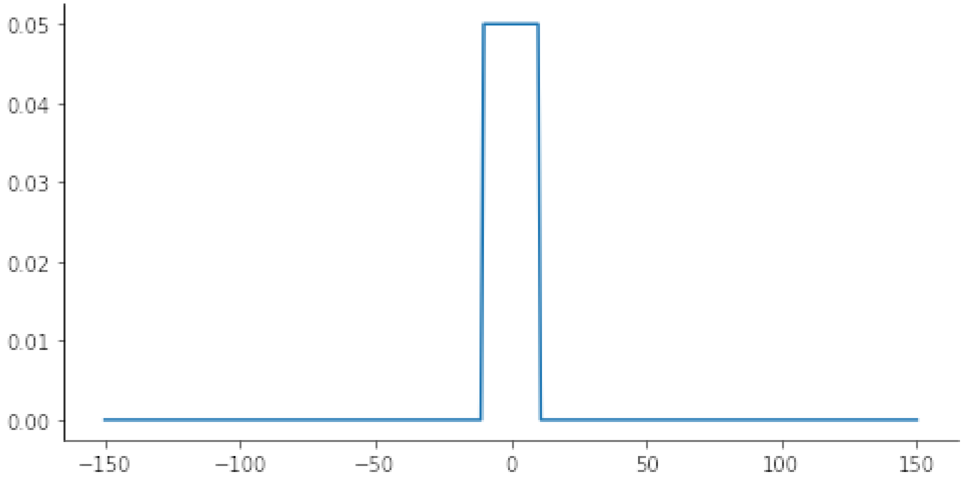
$$(f * g)(t) = \sum_{k=-\infty}^{\infty} f(t-k)\, g(k)$$



Convolutions - Examples

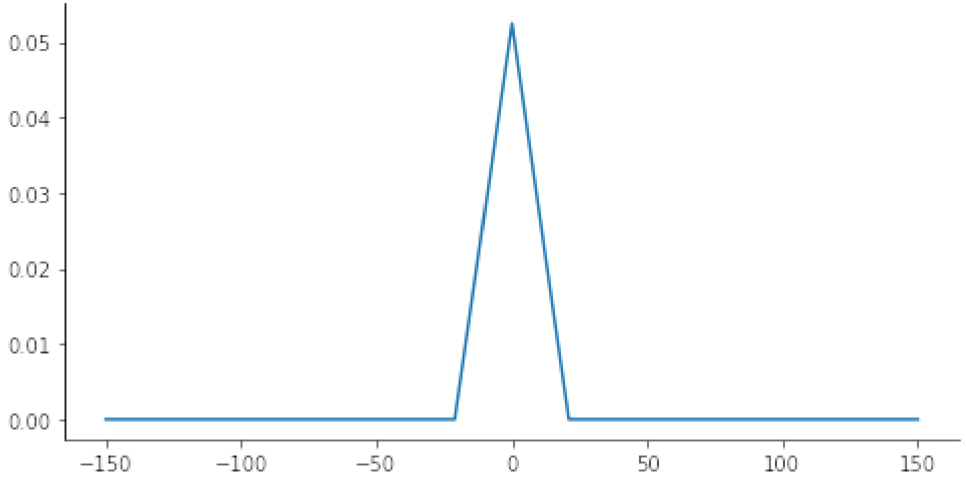


$*$

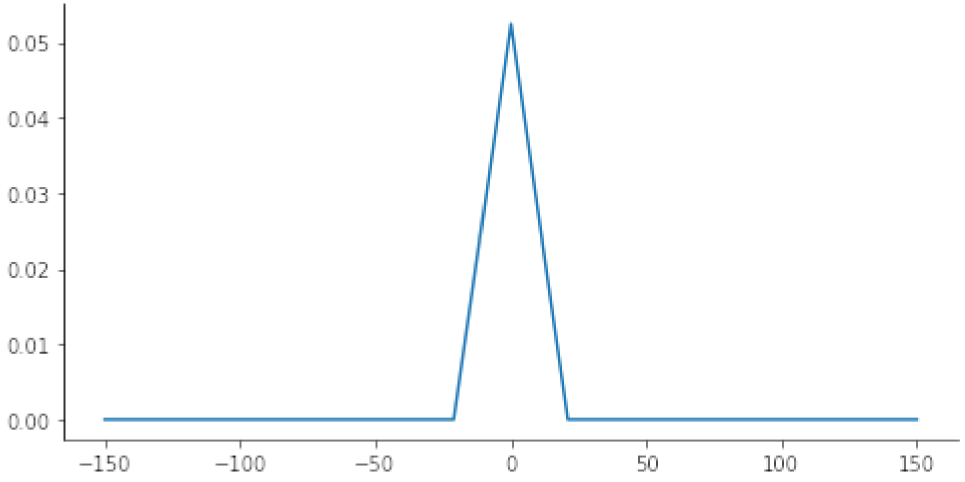


$=$

$$(f * g)(t) = \sum_{k=-\infty}^{\infty} f(t-k)\, g(k)$$



Convolutions - Examples

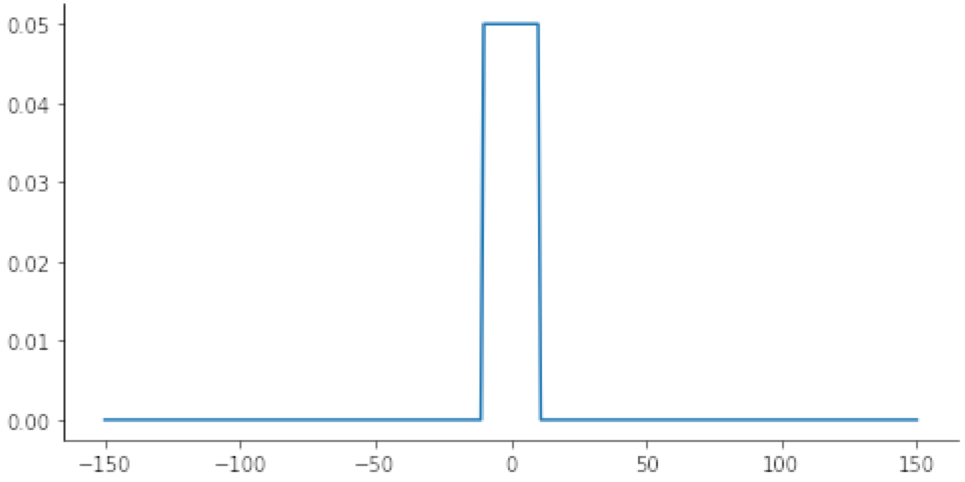


$$(f * g)(t) = \sum_{k=-\infty}^{\infty} f(t-k)\, g(k)$$

$*$

$=$

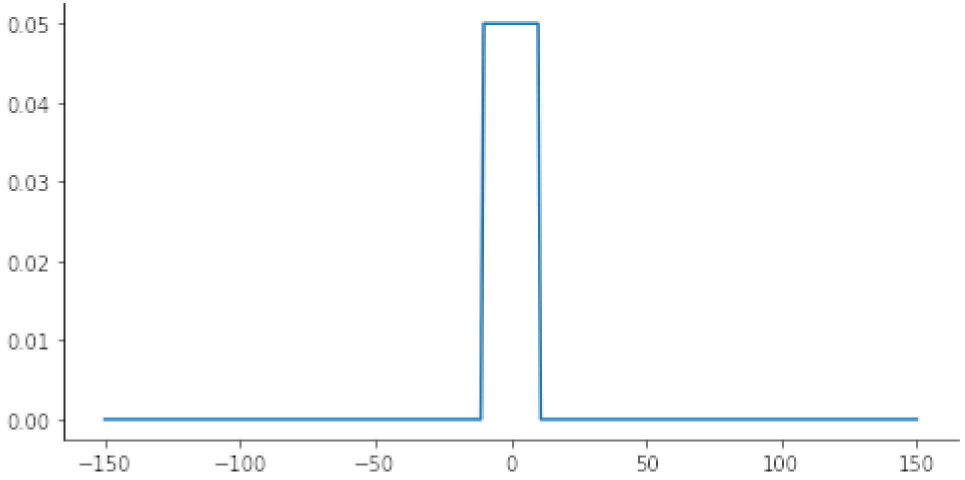
?



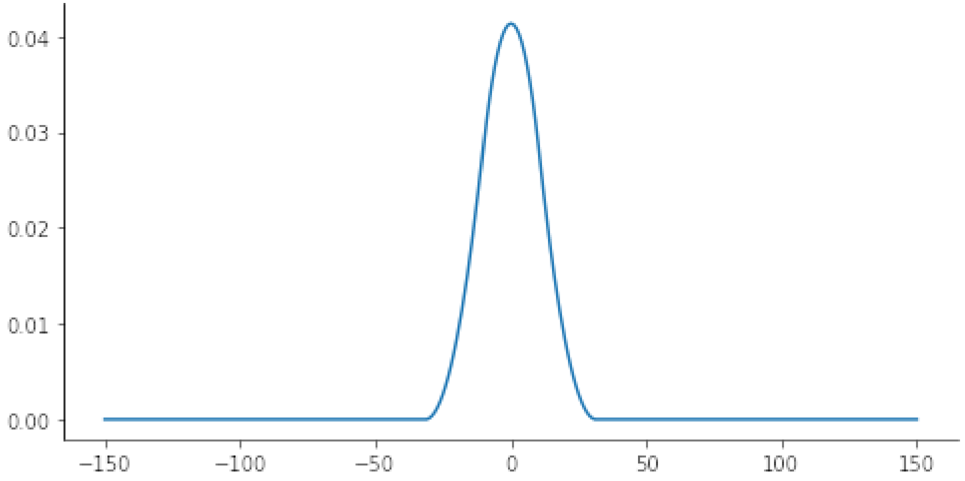
Convolutions - Examples

$$(f * g)(t) = \sum_{k=-\infty}^{\infty} f(t-k)\, g(k)$$

3



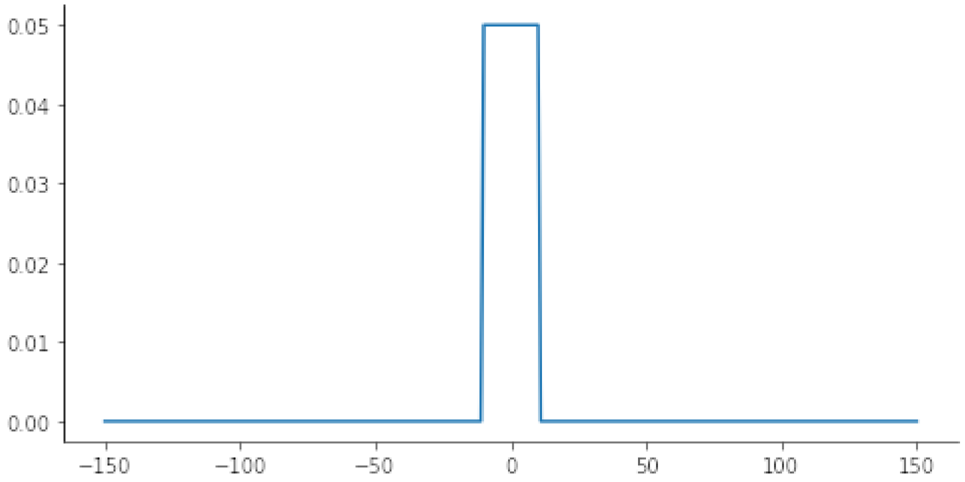
=



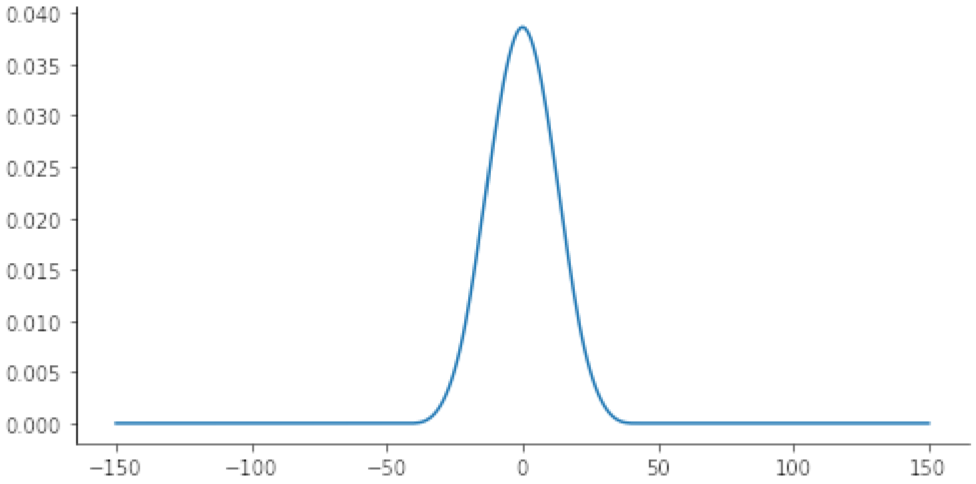
Convolutions - Examples

$$(f * g)(t) = \sum_{k=-\infty}^{\infty} f(t-k)\, g(k)$$

4



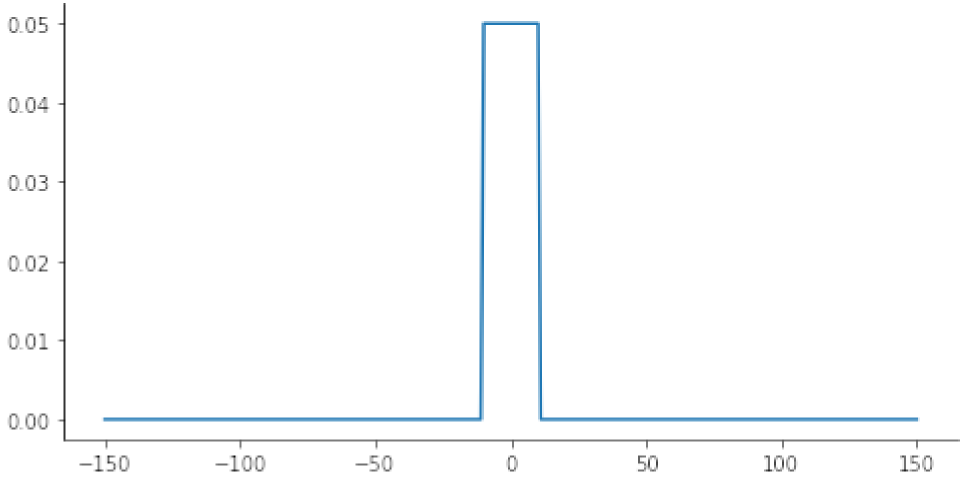
=



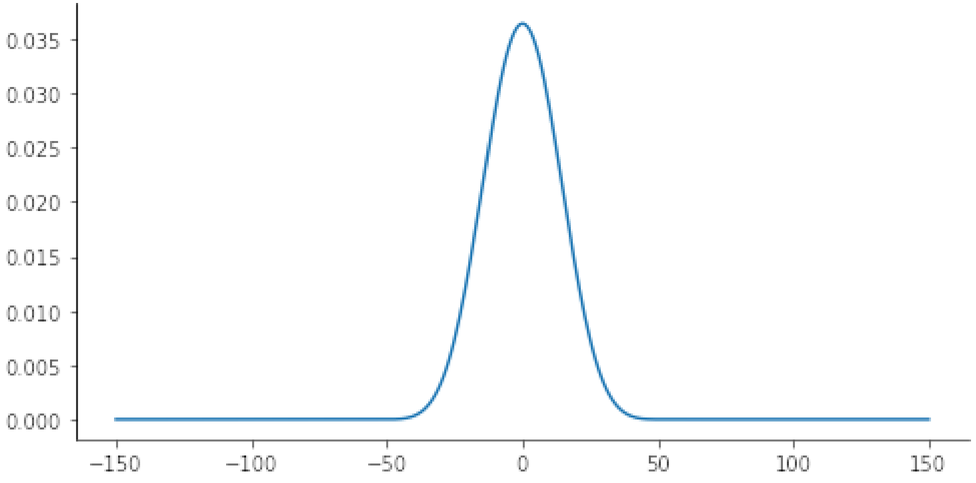
Convolutions - Examples

$$(f * g)(t) = \sum_{k=-\infty}^{\infty} f(t-k)\, g(k)$$

5



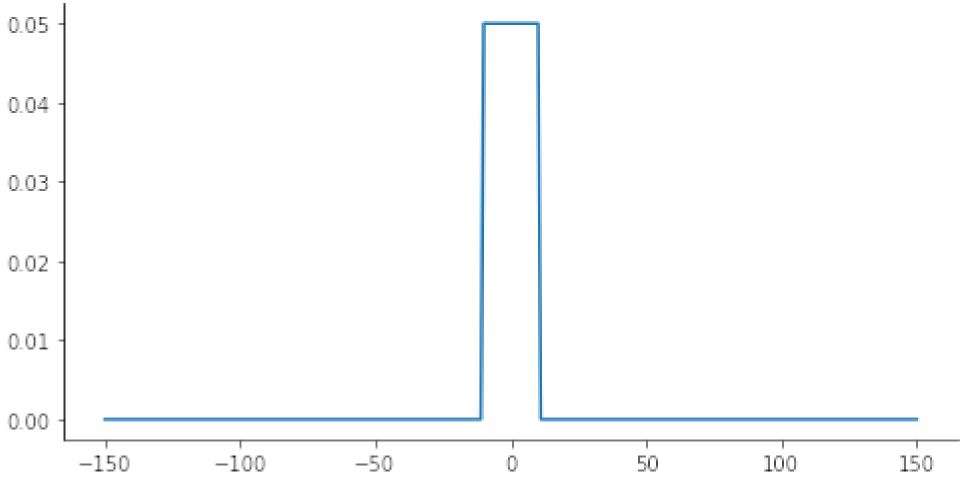
=



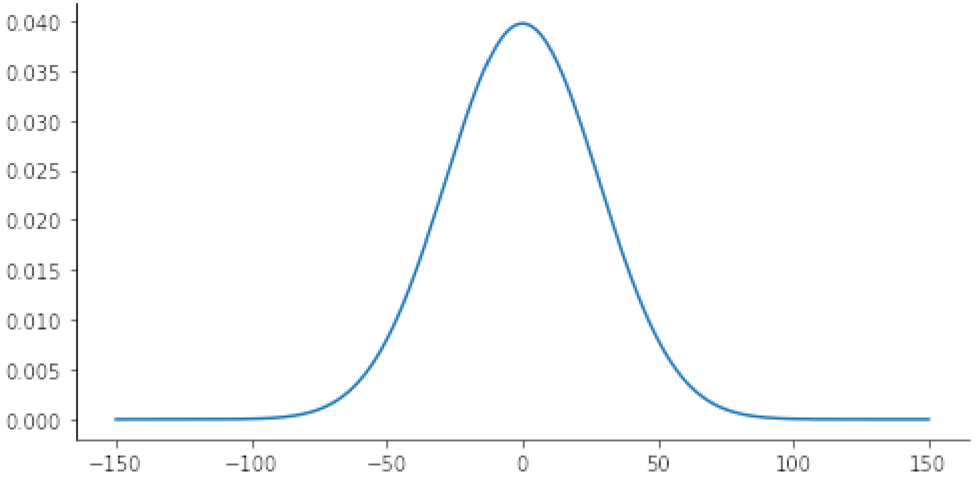
Convolutions - Examples

$$(f * g)(t) = \sum_{k=-\infty}^{\infty} f(t-k)\, g(k)$$

20



=



Remember Associativity

$(f * g) * h = f * (g * h)$

$f * g_{\text{Gaussian}} \approx f * g_{\text{box}} * g_{\text{box}} * \cdots * g_{\text{box}}$

Very efficient

Summary

- Rolling Average:
  - Simple
  - Very fast/time-efficient
  - Unexpected behaviour for some frequencies (flipping)
- Weighted Average:
  - Generalisation of moving average
- Gaussian smoothing
  - Weighted averaging with Gaussian Kernel
  - No flipping behaviour
  - Less time-efficient
- Convolutions:
  - Mathematical description of weighted averaging
  - Rolling average and Gaussian smoothing are special cases