## Time Series Encoding (encode)

This example shows how to transform a time series into fixed-size windows and train an autoencoder to learn a compact latent representation (p -> k) of these windows.

Prerequisites - R packages: daltoolbox, ggplot2 - Python with PyTorch accessible via reticulate (the backend is loaded by internal functions)

# Loading required packages  
library(daltoolbox)

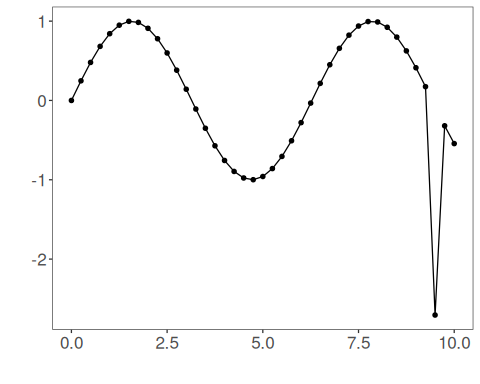
## Series for study

data(tsd)  
tsd$y[39] <- tsd$y[39] \* 6 # inject a synthetic outlier for illustration in the plot

sw\_size <- 5 # sliding window size (p)  
ts <- ts\_data(tsd$y, sw\_size) # convert the series into windows with p columns  
ts\_head(ts, 3) # view the first 3 windows

## t4 t3 t2 t1 t0  
## [1,] 0.0000000 0.2474040 0.4794255 0.6816388 0.8414710  
## [2,] 0.2474040 0.4794255 0.6816388 0.8414710 0.9489846  
## [3,] 0.4794255 0.6816388 0.8414710 0.9489846 0.9974950

library(ggplot2)  
plot\_ts(x = tsd$x, y = tsd$y) + # series plot with the outlier peak  
 theme(text = element\_text(size = 16))



## Data sampling

samp <- ts\_sample(ts, test\_size = 5) # hold out the last 5 windows for test  
train <- as.data.frame(samp$train)  
test <- as.data.frame(samp$test)

## Train the model

auto <- autoenc\_e(5, 3) # reduce from 5 -> 3 dimensions (p -> k)  
auto <- fit(auto, train)

## Encoding evaluation (train)

print(head(train)) # original windows (p columns)

## t4 t3 t2 t1 t0  
## 1 0.0000000 0.2474040 0.4794255 0.6816388 0.8414710  
## 2 0.2474040 0.4794255 0.6816388 0.8414710 0.9489846  
## 3 0.4794255 0.6816388 0.8414710 0.9489846 0.9974950  
## 4 0.6816388 0.8414710 0.9489846 0.9974950 0.9839859  
## 5 0.8414710 0.9489846 0.9974950 0.9839859 0.9092974  
## 6 0.9489846 0.9974950 0.9839859 0.9092974 0.7780732

result <- transform(auto, train) # encodings (k columns)  
print(head(result))

## [,1] [,2] [,3]  
## [1,] -0.05195337 -0.2077751 -0.7935887  
## [2,] 0.14650320 -0.4408252 -0.9251222  
## [3,] 0.34201825 -0.6656633 -0.9977869  
## [4,] 0.50701660 -0.8640385 -0.9966494  
## [5,] 0.62439787 -1.0200856 -0.9276072  
## [6,] 0.67961514 -1.1192963 -0.7980617

## Encoding of the test set

print(head(test))

## t4 t3 t2 t1 t0  
## 1 0.9893582 0.9226042 0.7984871 0.6247240 0.4121185  
## 2 0.9226042 0.7984871 0.6247240 0.4121185 0.1738895  
## 3 0.7984871 0.6247240 0.4121185 0.1738895 -2.7054403  
## 4 0.6247240 0.4121185 0.1738895 -2.7054403 -0.3195192  
## 5 0.4121185 0.1738895 -2.7054403 -0.3195192 -0.5440211

result <- transform(auto, test)  
print(head(result))

## [,1] [,2] [,3]  
## [1,] 0.61565745 -1.13262951 -0.4319700  
## [2,] 0.50173873 -1.06238270 -0.1845812  
## [3,] 0.02738531 -1.98244941 1.7437426  
## [4,] -0.82550502 0.05612845 1.6775781  
## [5,] -0.81720978 -0.30351767 1.3268173