

Kernels for sequentially ordered data
Dr. Franz Kiraly
(University College London)

May 4, 2016

Kernel learning is a general framework providing methodology for descriptive/exploratory statistics, non-linear regression and classification learning, for objects of any kind, for example time series, text, matrices, vectors up to mirror symmetries, and so on - given the right feature representation encoded by the so-called kernel, a non-linear scalar product.

After briefly reviewing the kernel learning framework and prior work on learning with structured objects or invariances, we present methodological foundations for dealing with sequential data of any kind, such as time series, sequences of graphs, or strings. Our approach is based on signature features which can be seen as an ordered variant of sample (cross-)moments; it allows to obtain a “sequentialized” version of any static kernel. The sequential kernels are efficiently computable for discrete sequences and are shown to approximate a continuous moment form in a sampling sense.

A number of known kernels for sequences arise as “sequentializations” of suitable static kernels: string kernels may be obtained as a special case, and alignment kernels are closely related up to a modification that resolves their open non-definiteness issue. Our experiments indicate that our signature-based sequential kernel framework may be a promising approach to learning with sequential data, such as time series, that allows to avoid extensive manual pre-processing. (Joint with Harald Oberhauser)