WHAT DOES BACK PROPAGATION COMPUTE?

EDOUARD PAUWELS

We are interested in nonsmooth analysis of backpropagation as implemented in modern machine learning librairies, such as Tensorflow or Pytorch. First I will illustrate how blind application of differential calculus to nonsmooth objects can be problematic, requiring a proper mathematical model. Then I will introduce a weak notion of generalized derivative, named conservativity, and illustrate how it complies with calculus and optimization for well structured objects. We provide stability results for empirical risk minimization similar as in the smooth setting for the combination of nonsmooth automatic differentiation, minibatch stochastic approximation and first order optimization. This is joint work with Jérôme Bolte.