SCIENCES & BIOENGINEERING SCIENCES

The Research Group

Artificial Intelligence

has the honor to invite you to the public defense of the PhD thesis of

Nixon Ronoh

to obtain the degree of Doctor of Sciences

Title of the PhD thesis:
Natural Gradient Based Adaptive Markov Chain Monte Carlo

Promotors:
Prof. dr. Ann Nowé (VUB)
Prof. dr. Bernard Manderick (VUB)
Prof. dr. Ambrose Kiprop (Moi University)

The defense will take place on
Tuesday, January 16, 2024 at 15h through a live stream:
https://web.facebook.com/events/1569070020331275/

Members of the jury

Prof. dr. Viviane Jonckers (VUB, chair)
Prof. dr. Beat Signer (VUB, secretary)
Prof. dr. Jonathan C-W Chan (VUB)
Prof. dr. Yvan Saeys (Universiteit Gent)
Prof. dr. Pedro Isasi (Carlos III of Madrid University)
Prof. dr. Rafael Bello (Universidad Central "Marta Abreu" de Las Villas)

Curriculum vitae

Nixon Ronoh obtained a B.Sc. in Applied Mathematics (2007), and an M.Phil. in Applied Mathematics (2010) from Moi University, Kenya. He is currently a lecturer at the department of Mathematics and Computing at the Moi University, since April 2012, specialising in Applied Mathematics including Numerical Analysis and Partial Differential Equations. Nixon was awarded a sandwich PhD scholarship in the VLIR-IUC program with Moi University, Kenya, to do a PhD in the AI Lab of Vrije Universiteit Brussel in the area of Bayesian Machine Learning.

Abstract of the PhD research

Adaptive Metropolis (AM) is the benchmark for adaptive Markov chain Monte Carlo (MCMC) sampling. AM estimates the covariance of the distribution to sample from, called the target, using samples generated so far. This is based on the fact that the optimal covariance of the Gaussian proposal distribution is proportional to the covariance of the target.

The hillclimber variant of Covariance Matrix Adaptation Evolution Strategies, (1+1)- CMAES, can be turned into a MCMC-sampler called MCMA. Experiments have shown that MCMA performs as well as AM. It uses another adaptation scheme: the parameters of the covariance of the proposal are adapted such that the probability of generating better candidates is improved in each iteration. This makes sense since the candidate generated by the proposal is always accepted when better than the current sample. This adaptation scheme comes down to gradient descent in the space of parameters of the proposal equipped with the Euclidean metric. Here, geodesics, i.e. shortest paths connecting two points, and straight lines coincide. However, the Euclidean metric between parameters of distributions is not an accurate measure of (dis)similarity between the distributions themselves.

Natural Evolution Strategies (NES) tackle this problem in a principled way. Fisher information is used as a non-Euclidean metric to structure the set of symmetric positive definite matrices as a Riemannian manifold. Adaptation takes place on this manifold and follows the direction of the natural gradient as opposed to the vanilla gradient used in CMAES and MCMA. Geodesics are curves on this manifold that are not straight anymore and the natural gradient is tangent to the geodesic of interest.

This adaptation scheme is invariant under affine transformations and makes the sampler insensitive to covariances present in the target.

In this research we consider 1+1 variants of exponential and separable NES that can be transformed in a straightforward way into MCMC samplers called MxNES and MsNES, respectively. Additionally, since adaptation as used in Ms/xNES does not guarantee convergence towards the target, we consider 1) stopping adaptation half way, and 2) diminishing adaptation at some predetermined rate. We compare performance using 6 measures on a test suite of 7 targets for state space dimensions ranging from 2 to 50.