

f -divergences and Monte Carlo methods

Summary - english version

Ph.D. candidate OLARIU Emanuel Florentin

Advisor Professor LUCHIAN Henri

This thesis broadly concerns the use of f -divergences mainly for variance reduction in Monte Carlo (MC) integration. A f -divergence is a particular type of measure for two probability distributions. MC is a classical randomized method for solving various types of problem for which we do not know analytical solutions; it is based on generating samples from particular distributions. Hence the problem of comparing distributions comes up naturally.

By the name of one of the first who studied them, these divergences are also called Csiszár divergences, and are generated by convex functions. More general a divergence measure is a function of two probability density (or distribution) functions, which has nonnegative values and becomes zero only when the two arguments (distributions) are the same. Often, a divergence is not a symmetric function but can be easily symmetrized.

There are many techniques for reducing the variance of the MC estimator and one of these is Importance Sampling (IS). Monte Carlo, f -divergences and various directions of variance minimization for IS and MC estimators are described in more detail in Chapter 2.

Monte Carlo methods, although used in the beginning for stochastic simulation only, covers today a wide range of problems which could benefit from randomness and adjacent prop-

erties. Generally speaking any technique which approaches a problem using a “large” number of random samples for various computations will take the famous name. This method is intended to solve problems for which deterministic/analytic approaches are not available, or give poor results.

We used MC method in two ways: for pricing financial derivatives known as options, and for estimation of rare-events probabilities. Both these applications are again linked with the use of f -divergences.

In Chapter 3 we developed some techniques for pricing two option styles. Spread European options are valued using IS and minimizing various divergences; this approach is compared with the least squares method for direct variance minimization. Bermudan options are priced using a modified version of MRAS algorithm, involving sampling importance resampling following the reference distributions from the standard algorithm.

There are many financial instruments for which closed form formulae cannot be derived from the existing mathematical models. One example of such model is the classical result of Black and Scholes which cover only a small part of the entire spectra of derivatives, especially for multivariate contingent claims.

The option is a derivative instrument - because do not depend directly on the price of an asset (commodities, stocks, currencies or financial indexes).

An option is a contract between two parts (a seller and a buyer) in which, one - the buyer - buys the right to engage in a transaction concerning the asset (at a future date), from the second - the seller. The buyer has the right, but not the obligation to fulfill the above transaction, while the seller has the obligation to engage the transaction if the first party agrees with that. Therefore an option contract can be exercised or not

at the convened moment(s) in time.

Depending on the transaction involved, there are two types of options: if the transaction gives the right to buy the asset(s) is called call option, while, if the transaction gives the right to sell the asset(s) is called put option.

Depending on the moment when the transaction can be exercised, there are two main styles of options: European - the option can be exercised only at the expiration (maturity) date and American - the contract can be exercised any time between the writing and the expiration date. In between those two reference types exist many others, like: Bermudan option - the buyer has the right to exercise at a designated number of times, Canary option - the buyer has the right to exercise at a designated number of times but not before a time period etc.

In the first section of Chapter 3 we used the already described divergences combined with the least squares method in order to approximate the price for call european spread options; Kulback-Leibler divergence gives the best results in terms of variance reduction.

The second section of this Chapter is motivated by the problem of pricing Bermudan option. We modified a method successfully used for solving optimization problem named Model Reference Adaptive Search which is a two stages procedure: first generate data samples using a specific random procedure (from distribution with known parameters), and secondly update the parameters for the random procedure using the data from previous step. The calculus of parameters in the second step often involves random variable expectation which are estimated by MC simulation.

We used here importance sampling in the form of sampling importance resampling: from first step generated samples we resample (with replacement) using a multinomial distribution having probabilities proportional with their importance ratios. In this way we give more importance to samples which shift

towards another distribution; the main distributions we chose for resampling are those of reference in the original algorithm.

Our algorithm performs almost twice as fast as the standard algorithm having same standard errors - this means that our method is a reliable and faster method.

The problem of estimation of rare events probabilities appears frequently in the analysis of performance of communications systems (e.g. the probability of failure of a network system). The IS problem for this estimation consists in the increase of the frequency of rare-events by changing (to a more important) distribution. We introduced a new algorithm for such an estimation based on Rényi divergence instead of Kullback-Leibler divergence (the cross-entropy method). This algorithm with its stochastic counterpart and a version for solving continuous optimization problems are presented in Chapter 4.

The general procedure we described does not involve any specific family of distributions, the only restriction is that the search space consists of product-form probability density functions. We discussed an algorithm for estimation of probability of rare events and a version for continuous optimization. The results of numerical experimentation with these algorithms carried in the last section support their performances.

Numerical experiments carried in this chapter show that the time simulation effort is reduced by roughly a factor of 350. Our estimates are more accurate for smaller probabilities and have smaller relative errors. Results allow us to compare our method involving Rényi divergence minimization with cross-entropy method. We can see that our algorithm is at least as good in terms of estimates, relative error and execution time; the estimates are in general better although the differences remain small. We can conclude that our method is a good alternative

for estimating small probabilities, and probably a procedure for tuning the parameter α will improve these results.

For optimization purposes in the experiments we used the family of multivariate normal distributions with independent components, and we tested our procedure on some known optimization problems (Griewangk's, Pinter's functions, Rosenbrock's valley etc). We compared our results with those obtained with the MRAS algorithm, and we can see that they are quite similar.

The last chapter concerns the means for measuring the similarity of time series (historical and synthetic data). Time series are common in various fields of science: medicine, multimedia, computational finance etc, and synthetic datasets are used in prediction and computer simulations.

The relevance of synthetic time series relies first on the data generator used, and, secondly, on the accuracy of the similarity measures involved in the process of comparing the original and the synthetic time series. The critical issue is to choose an appropriate "distance" to represent the similarity between two time series. We proposed to use as similarity measure some known symmetrised divergences; our experiments prove that such measures combined with mean similarity give more accurate results in terms of trajectory quality. Numerical experiments show that we can measure with great accuracy by using simple instruments like mean similarity and symmetrized divergences; these tools are easier to compute than the usual features which include common statistics, extremal points, slope and filtered data.