An Application of Genetic Programming To Investment System Optimization

Charles E. Smith
Ned Davis Research, Inc.
600 Bird Bay Drive, West
Venice, Florida, 34292
Charlie@ndr.com

ABSTRACT

The Ned Davis Research Evolutionary Optimizer (NEO) is a tool that performs genetic programming optimization of quantitative investment systems. It selects combinations of commands, arguments, and parameters that form nearly optimal programs for investment timing or asset allocation. The primary inputs to NEO are a set of market or economic data series and a template that limits the types of code that can be generated. The output is a program (in NDR's proprietary language Technalyzer) that provides investment timing signals or asset allocation recommendations.

1. Introduction

NEO is an efficient generator of optimized software for investment systems. It includes numerous built-in functions for representing fitness, such as gain per annum, percent of possible market profit actually achieved, and other measures and combinations of measures. Also included are numerous built-in penalty functions to enforce boundary conditions, such as trades per year. NEO also allows the user to write custom fitness and penalty functions. Genetic search parameters such as population size can be set by the user or allowed to take default values. NEO is actually a hybrid optimizer that first uses evolutionary techniques to sample the search space and then, optionally, perform local hill-climbing on the best solution found after a preset number of generations.

2. Application Example

A real-world application of NEO involves regression analysis of some money-supply factors as predictors of period-to-period change in the Standard and Poor's 500 Stock Index. Included were such economic data series as gross domestic product, M3 and MZM money supply series, consumer price index, and the NDR commodity futures index. This process includes a search for the transformation functions that make each independent variable most predictive of the dependent variable, as well as a search for the individual regression coefficients. Such transformations as the Stochastic %K function are made available in the template code. The time periods to use for measuring rates of change of the individual economic measures are also being selected. The lengths of time to "lag" the measures (slide each independent-variable data series with respect to the dependent-variable data series) are also being optimized. The optimization process even finds the length of the period-to-period change for the dependent variable that is most effectively predicted. In this example, fitness was calculated as the index of determination, R^2. This approach produced monthly percentage change predictions for the S&P 500 (4.5 months forward) with an in-sample index of determination of 0.481 and excess returns of three percent. The search space for this problem is around $5x10^{15}$ combinations, each involving nineteen years of monthly data.

3. Other Features / Summary

Many useful genetic processing options have been included in NEO. For example, an exhaustive populate option provides for inclusion of every value in each specified template-argument range at least once in the initial population. Another option, "shuffle cross" provides that if the fitness of either offspring of a crossover operation is higher than the most-fit parent, then both offspring replace the parents. This maintains all of the genetic material of the original population (but tends to concentrate the search around fewer fit individuals).

A job-control facility runs multiple independent NEO optimization jobs in parallel on multiple host processors, keeping both slower and faster processors busy until all jobs in a list have been run.

Another application of NEO has been the "training" of artificial neural networks. Connection weights, biases, and transfer functions are programmed in the template file. Then NEO searches for optimal weights and biases.

NEO is a powerful tool for generating and optimizing quantitative investment systems.