Hyper-parameter optimization for strategy calibration

Yacine Debbabi

November 28, 2020

Contents

1 Objectives

Let us denote by $G(X;h)$ the (random) profit obtained from running a trading strategy with (random) market data X and using the strategy hyperparameter h. We assume $G(X;h) \sim N(\mu(h), \sigma^2(h))$. Our objective is to choose h to maximize a risk-penalized gain, i.e.

$$\hat{h} = \mathrm{argmax}_h \frac{\mu(h)}{\sigma(h)^\beta}, \tag{1}$$

where $\beta \in [0, 1/2]$ is a coefficient controlling the amount of risk penalization. Note that the objective function is not directly observable for a given h. The function value must be estimated by sampling several realizations for X. This yields noisy observations.

Our objective is to discuss (a) how to select an optimum over a finite set of hyperparameter values, and (b) how to efficiently search the parameter space.

2 Optimization on finite parameter sample set

We estimate the objective function for a given hyperparameter value h as $\hat{\mu}_n(h)/\hat{\sigma}_n^\beta(h)$, where

$$\hat{\mu}_n(h) = \frac{1}{n}\sum_{i=1}^{n} G(X_i;h), \tag{2}$$

$$\hat{\sigma}_n^2(h) = \frac{1}{n-1}\sum_{i=1}^{n}(G(X_i;h) - \hat{\mu}_n(h))^2, \tag{3}$$

and $X_1, ..., X_n$ are drawn from the distribution of X.

3 An explanatory model

The model assumes a response function follows a binomial distribution, i.e. $f(X;h) \sim B(p(h))$. Say we observe n realizations of that random variable on a grid of hyperparameter values $h_1, ..., h_p$. We can estimate $\hat{p}(h) = \sum_1^n f(X_i;h)$. We try to

$$\hat{h} = \text{argmax}_h \mathbb{E}\left[f(X;h)\right] / \text{Var}\left[f(X;h)\right]^{\beta/2} \tag{4}$$

4 A need for robust objective functions

Optimization problems relying on a data-dependent objective function might be prone to producing spurrious results. Examples of possible issues include:

- Objective function values can be non robust against minor changes in the underlying dataset,
- The objective function might exhibit large discontinuities that are data dependent.

How can we guarantee robustness of the results?

Interesting links:

- Robust statistics