Online Portfolio Hedging with the Weak Aggregating Algorithm

Najim Al-baghdadi[1] Yuri Kalnishkan[1] David Lindsay[2] Siân Lindsay[2]

[1]Department of Computer Science, Royal Holloway, University of London, Egham, United Kingdom

[2]AlgoLabs, Bracknell, United Kingdom

COPA, Brighton, August 2022

The Problem

- Hedging is the act of protecting an investment against unfavorable moves in the market by trading a negatively correlated asset or investment instrument.
- Here we explore using prediction with expert advice algorithms to find optimal hedge decisions given a pool of hedging strategies.

Outline of talk

1. Cylinder Hedging Model
2. Weak Aggregating Algorithm
3. WAA for Hedging
4. Empirical Results
5. Conclusions

1 Cylinder Hedging Model

2 Weak Aggregating Algorithm

3 WAA for Hedging

4 Empirical Results

5 Conclusions

Cylinder Hedging Model

- We will be looking at a special case of a **Financial Market Makers (MMs)** hedging strategy
- The model has two main parameters:
  1. A pair of long and short limits (typically specified in US dollars)
  2. A hedge fraction specifying how much to hedge

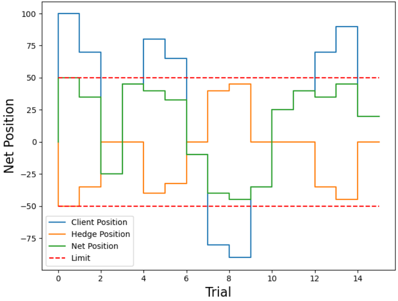


Figure: Cylinder Model Position

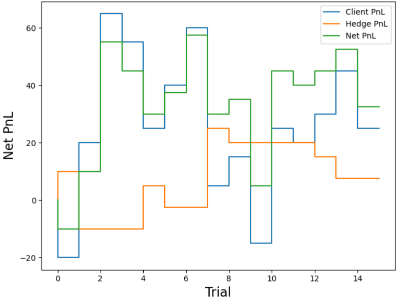


Figure: Cylinder Model PnL

Cylinder Hedging Model

- Directional indicators can improve on the models ability to hedge effectively

Algorithm 1 Cylinder Hedging Model

Parameters: long/short Limit, Hedge fraction and Skew: $L_l, L_s, H_l, H_s, S_l, S_s$
Directional indicators Id_t, $t = 1, 2, \ldots$

for $t = 1, 2, \ldots$ **do**
  if $Position_t^C > L_l + (L_l \times S_l \times Id_t)$ **then**
    Hedge $\text{Fraction}_t \leftarrow H_l$
  end
  if $Position_t^C < L_s + (L_s \times S_s \times Id_t)$ **then**
    Hedge $\text{Fraction}_t \leftarrow H_s$
  end
  else
    Hedge $\text{Fraction}_t \leftarrow 0$
  end
end

Cylinder Hedging Model

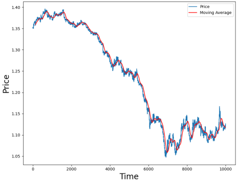


Figure: Price of Underlying Asset

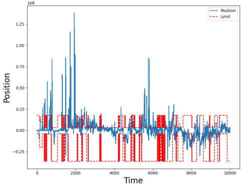


Figure: Client Position

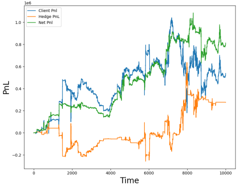


Figure: Client, Hedge and Net PnL

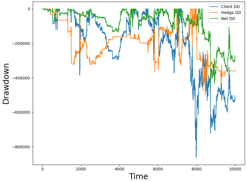


Figure: Client, Hedge and Net Drawdown

1 Cylinder Hedging Model

2 Weak Aggregating Algorithm

3 WAA for Hedging

4 Empirical Results

5 Conclusions

Prediction with Expert Advice Framework

- On every step $t = 1, 2, \ldots$, the learner L produces a *prediction* $\gamma_t \in \Gamma$, where Γ is a known prediction space.
- The nature produces a *loss function* $\lambda_t : \Gamma \to \mathbb{R}$ and the learner suffers loss $\ell_t = \lambda_t(\gamma_t)$.
- We measure the performance of L by the *cumulative loss* over T steps given by

$$\mathrm{Loss}_T(\mathcal{L}) = \sum_{t=1}^{T} \ell_t \quad .$$

Prediction with Expert Advice Framework

- Suppose that there are N *experts* $\mathcal{E}_n$, $n = 1, 2, \ldots, N$, making prediction in the same environment as $\mathcal{L}$.
- We want the cumulative loss $\mathrm{Loss}_T(\mathcal{L})$ to be small compared to the minimum of experts' losses $\mathrm{Loss}_T(\mathcal{E}_n) = \sum_{t=1}^{T} \ell_t^n$.

Protocol 1 Prediction with Expert Advice Protocol

for $t = 1, 2, \ldots$ **do**
  experts $\mathcal{E}_n$ output predictions $\gamma_t^n \in \Gamma$, $n = 1, 2, \ldots, N$
  learner $\mathcal{L}$ outputs a prediction $\gamma_t \in \Gamma$
  nature produces a function $\lambda_t : \Gamma \to \mathbb{R}$
  experts $\mathcal{E}_n$ suffer losses $\ell_t^n = \lambda_t(\gamma_t^n)$, $n = 1, 2, \ldots, N$
  learner $\mathcal{L}$ suffers loss $\ell_t = \lambda_t(\gamma_t)$
end

Weak Aggregating Algorithm

- Let Γ be a convex set so that for any $\gamma_1, \gamma_2, \ldots, \gamma_N \in \Gamma$ and probabilities $p_1, p_2, \ldots, p_N$ ($p_n \geq 0$ for $n = 1, 2, \ldots, N$ and $\sum_{n=1}^N p_n = 1$) the convex combination $\gamma = \sum_{n=1}^N p_n \gamma_n$ is defined and belongs to Γ.
- In order to obtain performance bounds for WAA, one needs to assume convexity of loss functions λ_t; this ensures the inequality $\ell_t \leq \sum_{n=1}^N p_{t-1}^n \ell_t^n$.
- We will also need losses to be bounded. Let $L \in \mathbb{R}$ be such that

$$\max_{n=1,2,\ldots,N} \ell_t^n - \min_{n=1,2,\ldots,N} \ell_t^n \leq L$$

Weak Aggregating Algorithm

- A learner following the WAA protocol with equal initial weights and a learning rate $\eta_t = c/\sqrt{t}$ where $c = 2\sqrt{\ln N}/L$, can ensure the following bound on loss:

$$\mathrm{Loss}_T(\mathcal{L}) \leq \mathrm{Loss}_T(\mathcal{E}_n) + L\sqrt{T \ln N}$$

for all $T = 1, 2, \ldots$ and all experts $\mathcal{E}_n$, $n = 1, 2, \ldots, N$.

Algorithm 2 Weak Aggregating Algorithm

Parameters: Initial distribution $q_1, q_2, \ldots, q_N$, $q_n \geq 0$ for $n = 1, 2, \ldots$ and $\sum_{n=1}^{N} = 1$
Learning rates $\eta_t > 0$, $t = 1, 2, \ldots$

let $L_0^n = 0$, $n = 1, 2, \ldots, N$
for $t = 1, 2, \ldots$ **do**
  calculate weights $w_{t-1}^n = q_n e^{-\eta_t L_{t-1}^n}$, $n = 1, 2, \ldots, N$
  normalise the weights $p_{t-1}^n = w_{t-1}^n / \sum_{i=1}^{N} w_{t-1}^i$, $n = 1, 2, \ldots, N$
  read experts' predictions $\gamma_t^n \in \Gamma$, $n = 1, 2, \ldots, N$
  output $\gamma_t = \sum_{n=1}^{N} p_{t-1}^n \gamma_t^n$
  read experts losses ℓ_t^n, $n = 1, 2, \ldots, N$
  update $L_t^n = L_{t-1}^n + \ell_t^n$, $n = 1, 2, \ldots, N$
end

Discounted Loss

- Suppose that we are given coefficients $\alpha_1, \alpha_2, \ldots \in (0,1]$. Let the cumulative discounted loss for a learner $\mathcal{L}$ be given by

$$\widetilde{\mathrm{Loss}}_T(\mathcal{L}) = \sum_{t=1}^{T} \lambda(\gamma_t) \left(\prod_{s=t}^{T-1} \alpha_s \right) = \alpha_{T-1} \widetilde{\mathrm{Loss}}_{T-1}(\mathcal{L}) + \lambda(\gamma_T) \;;$$

- If L is known in advance and all discounting factors are equal and less than 1, one can take

$$\eta_t = \eta = \frac{2\sqrt{2(1-\alpha)\ln N}}{L}$$

and ensure for equal weights $q_1 = q_2 = \ldots = q_N = 1/N$ the bound

$$\mathrm{Loss}_T(\mathcal{L}) \leq \mathrm{Loss}_T(\mathcal{E}_n) + L\sqrt{\frac{\ln N}{2(1-\alpha)}} \tag{1}$$

for all $T = 1, 2, \ldots$ and all experts $\mathcal{E}_n$, $n = 1, 2, \ldots, N$.

Discounted Loss

Algorithm 3 Weak Aggregating Algorithm with Discounting

Parameters: Initial distribution $q_1, q_2, \dots, q_N$, $q_n \geq 0$ for $n = 1, 2, \dots$ and $\sum_{n=1}^{N} = 1$.
Discounting factors $\alpha_1, \alpha_2, \dots \in (0, 1]$.
Learning rates $\eta_t > 0$, $t = 1, 2, \dots$

let $L_0^n = 0$, $n = 1, 2, \dots, N$
for $t = 1, 2, \dots$ **do**
  calculate weights $w_{t-1}^n = q_n e^{-\eta_t \alpha_{t-1} L_{t-1}^n}$, $n = 1, 2, \dots, N$
  normalise the weights $p_{t-1}^n = w_{t-1}^n / \sum_{i=1}^{N} w_{t-1}^i$, $n = 1, 2, \dots, N$
  read experts' predictions $\gamma_t^n \in \Gamma$, $n = 1, 2, \dots, N$
  output $\gamma_t = \sum_{n=1}^{N} p_{t-1}^n \gamma_t^n$
  read experts losses ℓ_t^n, $n = 1, 2, \dots, N$
  update $L_t^n = \alpha_{t-1} L_{t-1}^n + \ell_t^n$, $n = 1, 2, \dots, N$
end

1 Cylinder Hedging Model

2 Weak Aggregating Algorithm

3 WAA for Hedging

4 Empirical Results

5 Conclusions

WAA for Hedging

- Our pool of experts are a set of cylinder models
- A hedge decision is represented by $\gamma \in [-1, 0]$, where $\gamma_t = -1$ implies hedging out the entire client position and $\gamma_t = 0$ corresponds to a decision not to hedge over trial t
- It is natural to define loss in terms of the the MM's PnL resulting from facilitating client orders
- As PnL represents the MM's gain, we need to take its inverse when defining the loss. We can therefore take the loss at time t to be $\lambda(\gamma_t) = -\text{PnL}_t \gamma_t$

WAA for Hedging Loss function

- Here we will take a similar approach considering the loss function with the coefficients $u \geq 0$ and $v \geq 0$:

$$\lambda(\gamma) = -\left(\frac{u}{u+v}\text{PnL}\gamma + \frac{v}{u+v}\min(\text{PnL}\gamma, 0)\right)$$

- This allows a the learner to adjust their risk appetite adding more focus on losses that profit and minimise drawdown

1 Cylinder Hedging Model

2 Weak Aggregating Algorithm

3 WAA for Hedging

4 Empirical Results

5 Conclusions

Data Set

- Real-world currency exchange price data and client order data based on the trading behaviour of individuals opening positions with an FX MM
- EUR/USD over a 41 month period (Feb 2014 - June 2017) represented in hourly epochs

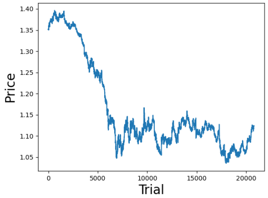


Figure: Price

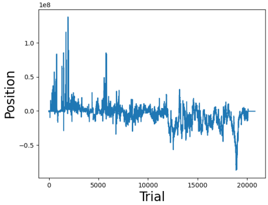


Figure: Position

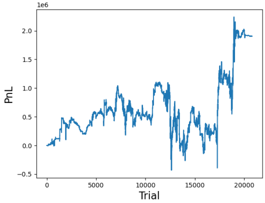


Figure: PnL

Experts

- Our pool of experts is the hedge fraction predictions from 100 unique cylinder models.

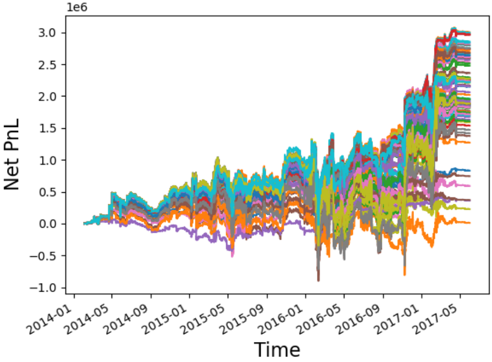
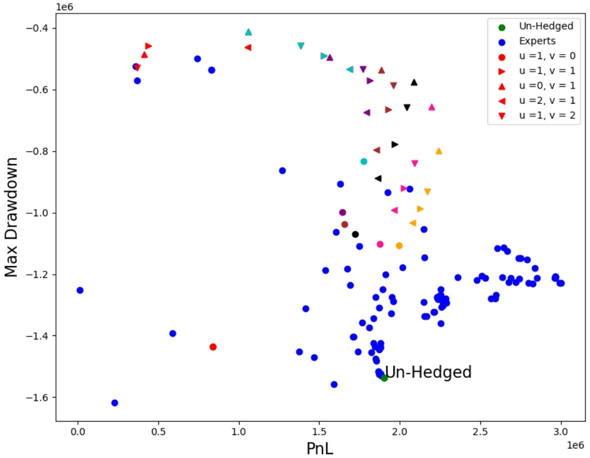


Figure: EUR/USD PnL

Results

EUR/USD PnL against Max Drawdown.**Discount Key** Red: 0%, Cyan: 2.5%, Purple: 5%, Black: 7.5%, Pink: 10%, Orange: 20%



1 Cylinder Hedging Model

2 Weak Aggregating Algorithm

3 WAA for Hedging

4 Empirical Results

5 Conclusions

Conclusions

- We have shown that the Weak Aggregating Algorithm (WAA) can be used to combine the predictions from a pool of cylinder hedging models to improve key performance metrics - namely the overall profit (PnL) - whilst simultaneously not compromising on the smoothness of returns by minimising drawdowns
- We have further introduced a method for applying discounted loss to the WAA