Online Learning for Stock Market Investment

Fall 2017, COMS 6998 Final Project

Kevin Wu
Gregory Johnsen
John Heintschel
Sasha Beltinova

Introduction

- Stock markets: a natural candidate for online learning algorithms.
    - Full feedback, immediately observable rewards, sequential structure

- Portfolio allocation problem: Given a set of stocks and an initial endowment of wealth, maximize cumulative rewards over a set time frame.

- Can an algorithm that minimizes hindsight regret for investment portfolios beat naive investment benchmarks in modern financial markets?

Portfolio Allocation: Approaches

- Markowitz (1952): “Modern Portfolio Theory”
  - Mean-variance optimization: parametric approach, need to reliably estimate μ, σ using historical prices
- Cover (1991): Universal Portfolios
  - First online learning approach to the problem
  - Invest uniformly in all constantly rebalanced portfolio strategies
  - Exponential complexity in the number of stocks
- Helmbold et al. (1998): Exponentiated Gradient
  - Linear in the number of stocks
  - Empirically, nearly just as good as Universal Portfolios on small number of stocks

Portfolio Allocation: Online Learning Setup

2.1 Portfolio Allocation Setup

The portfolio allocation problem can be formalized as follows.

- Given a set of K stocks, an agent invests wealth according to the non-negative weight vector $\boldsymbol{w} = (w^1, \ldots w^K)$ where $\sum_{i=1}^{K} w^i = 1$.
- An agent's total wealth at time t given a set of relative price changes[1] $\boldsymbol{x}_t = (x_t^1, \ldots x_t^K)$ is represented by the dot product $\boldsymbol{w} \cdot \boldsymbol{x_t}$.
- Over a fixed time period T, given a sequence of portfolio allocations $\boldsymbol{w}_1, \boldsymbol{w}_2, \ldots \boldsymbol{w}_T$ and a sequence of relative price changes $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots \boldsymbol{x}_T$, the final wealth, or the *cumulative reward*, is calculated as:

$$S(\cdot) = \prod_{t=1}^{T} (\boldsymbol{w}_t \cdot \boldsymbol{x}_t)$$

2.2 Constantly Rebalanced Portfolio, Best CRP

A constantly rebalanced portfolio (CRP) is an allocation method in which $\boldsymbol{w}_1 = \boldsymbol{w}_2 = \ldots \boldsymbol{w}_T$.[2] The best Constantly Rebalanced Portfolio, $\boldsymbol{w}^*$, is defined *in hindsight* by:

$$\boldsymbol{w}^* = \operatorname*{argmax}_{w} S(\cdot) = \operatorname*{argmax}_{w} \prod_{t=1}^{T} (\boldsymbol{w} \cdot \boldsymbol{x}_t)$$

Exponentiated Gradient

- Objective function: Maximize "expected" rewards while maintaining incremental updates to the weight vector *w*

$$\operatorname*{argmax}_{\boldsymbol{w}_{t+1}} \eta \log(\boldsymbol{w}_{t+1} \cdot \boldsymbol{x_t}) - D_{KL}(\boldsymbol{w}_{t+1} || \boldsymbol{w}_t)$$

- Weight update formula:

$$w_{t+1}^i = g(\boldsymbol{w}_t, \boldsymbol{x}_t) = \frac{w_t^i \exp(\eta x_t^i / \boldsymbol{w}_t \cdot \boldsymbol{x}_t)}{\sum_{j=1}^N w_t^j \exp(\eta x_t^j / \boldsymbol{w}_t \cdot \boldsymbol{x}_t)} \quad (1)$$

- Similarities to Hedge?

Exponentiated Gradient

Algorithm 1: Exponentiated Gradient

Parameter: η

Given K stocks and initial wealth S_0, initialize weight vector $\boldsymbol{w} = (1/K, ...1/K)$.

for $t \in \{1, \ldots, T\}$ **do**

  for $i \in \{1, \ldots, K\}$ **do**

    Observe reward x_i^t for stock

    Update weight $w_i^{t+1} = g(\eta, w_i^t, x_i^t)$. (See Equation 1)

  end for

  Update wealth. $S_t = \boldsymbol{w^t} \cdot \boldsymbol{x^t}$.

  Reallocate wealth according to $\boldsymbol{w_{t+1}}$.

end for

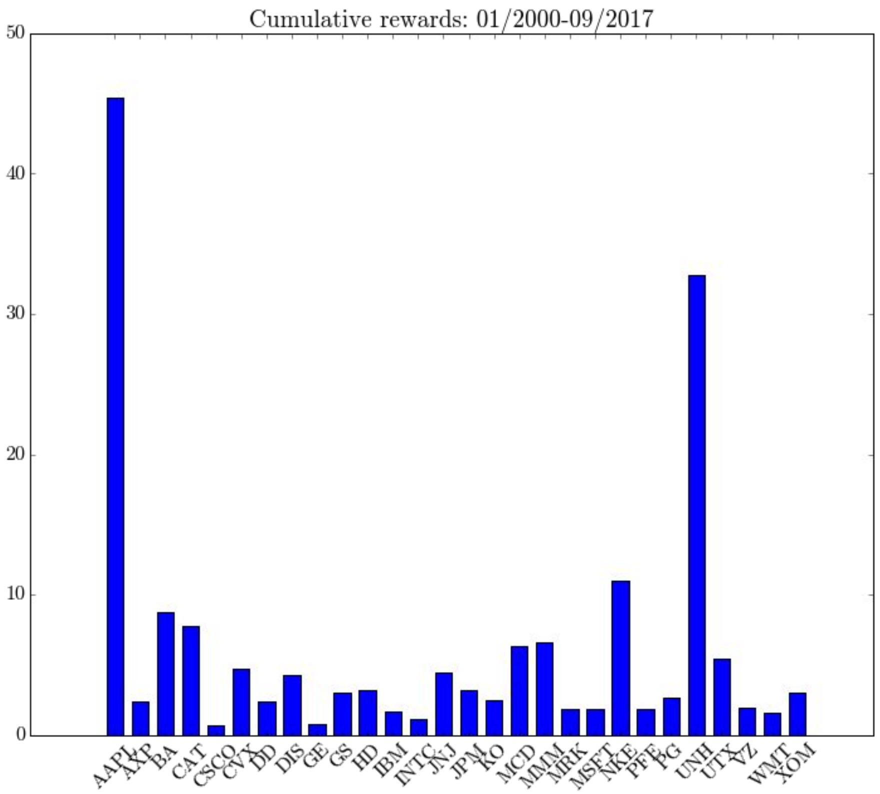
Experiments on U.S. Equities

- Used daily adjusted close prices of current Dow Jones Industrial Average (DJIA) components from 2000 to 2017.

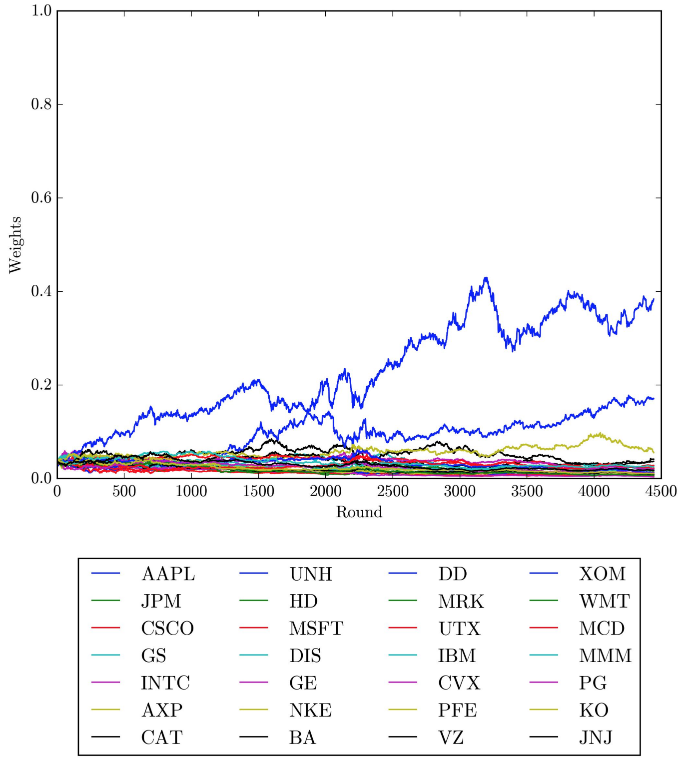
- Benchmarks:
  - Equal-Weighted Buy and Hold
    - Should beat this!
  - Best Stock
  - Best Constant Rebalanced Portfolio (CRP)

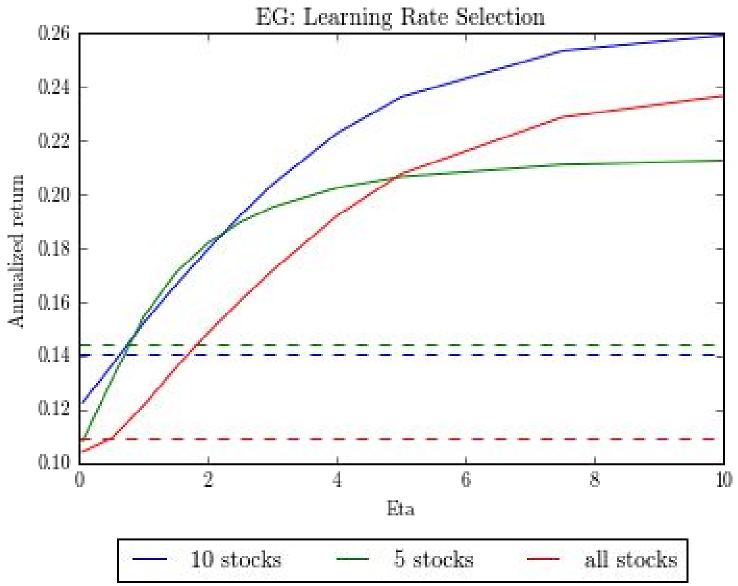
Experiments on U.S. Equities

AAPL	JPM
AXP	KO
BA	MCD
CAT	MMM
CSCO	MRK
CVX	MSFT
DD	NKE
DIS	PFE
GE	PG
GS	UNH
HD	UTX
IBM	VZ
INTC	WMT
JNJ	XOM

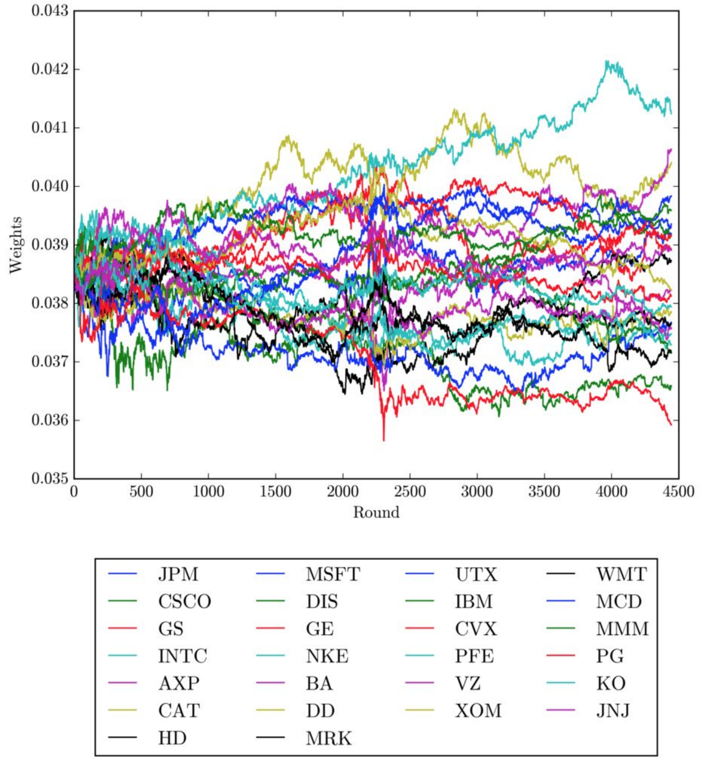


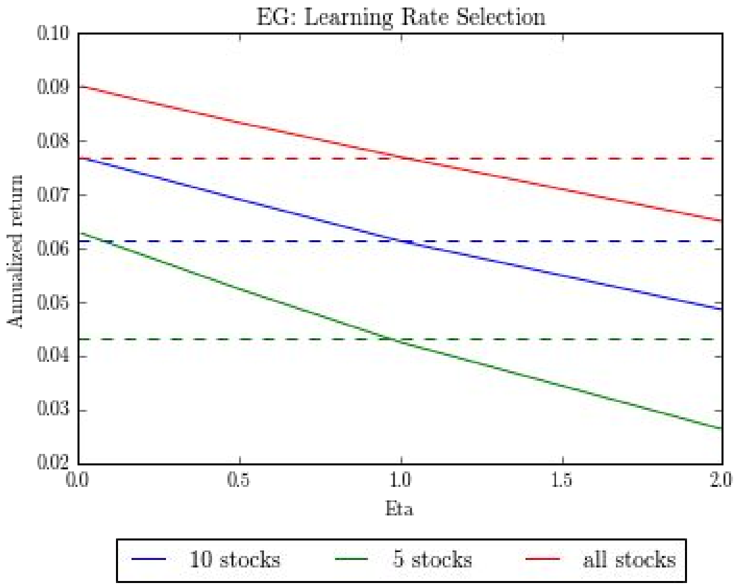
Portfolio Weights: “Winner-Take-All”





Portfolio Weights: “Winner-Take-All”





EG: “Sophisticated Experts”

- Same as regular EG, except portfolio is allocated among experts which may choose to either buy the stock or not.

Algorithm 2: Exponentiated Gradient with Sophisticated Experts

Parameters: η
Given K experts and initial wealth S_0, initialize weight vector $\boldsymbol{w} = (1/K, ...1/K)$.
for $t \in \{1, \ldots, T\}$ **do**
  for $i \in \{1, \ldots, K\}$ **do**
    Observe reward x_i^t for expert.
    Update weight for expert $w_i^{t+1} = g(\eta, w_i^t, x_i^t)$. (See Equation 1)
    Expert calls subroutine $EXP(x_i)$ to determine whether to BUY or NOT BUY.
  end for
  Update wealth. $S_t = \boldsymbol{w^t} \cdot \boldsymbol{x^t}$.
  Reallocate wealth according to $\boldsymbol{w_{t+1}}$.
end for

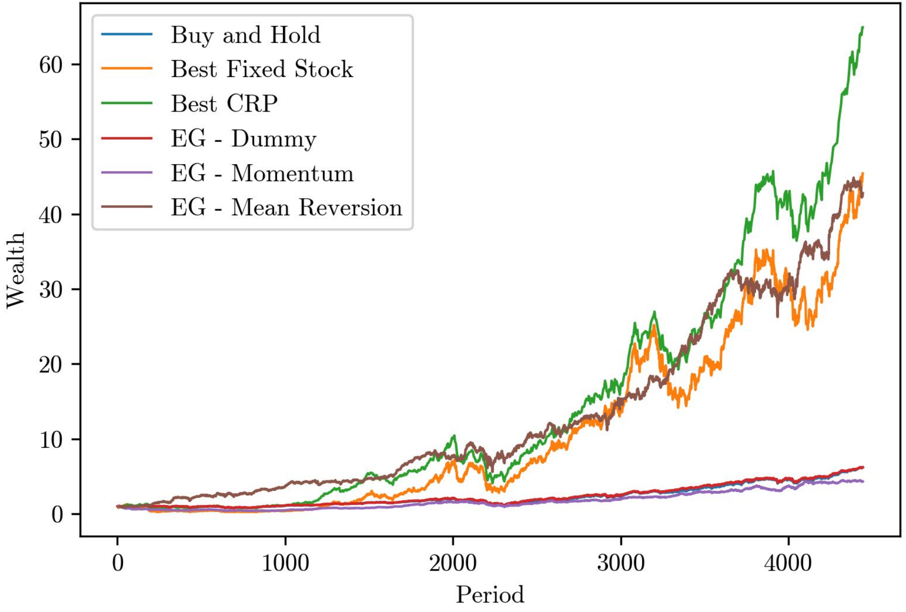
- Full allocation: if an expert decides not to buy, its allocation is redistributed amongst the “buying” experts in that period.

Expert Subroutines: Mean Reversion/Momentum

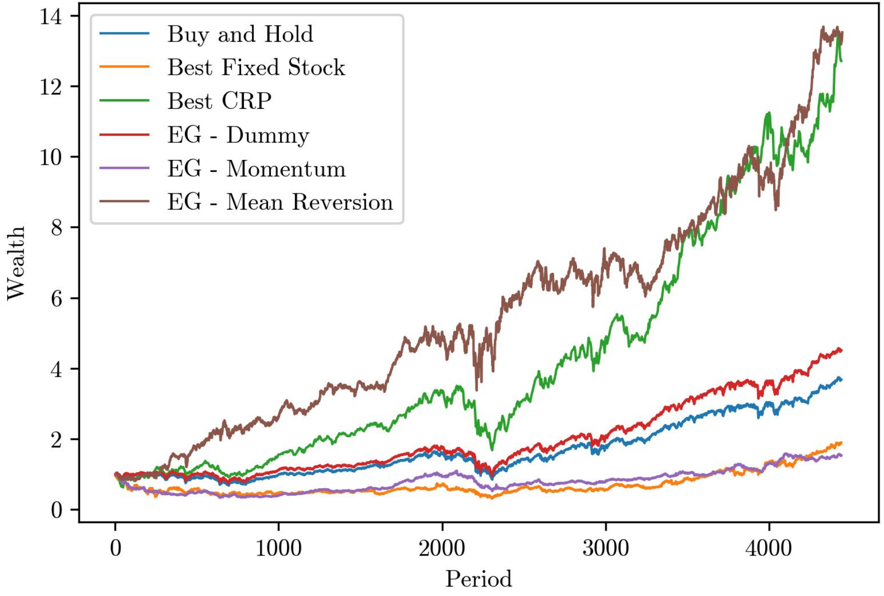
Algorithm 3: Moving Average Expert Subroutine: Mean Reversion

Parameters: n (window size), τ (moving average threshold)
Initialize expert for a given stock i.
Initialize a time-ordered list of at most n prices for i, $\boldsymbol{x'} = \emptyset$.
for $t \in \{1, \ldots, T\}$ **do**
  if $t < n$ **then**
    Add x_t^i to list of most recent prices $\boldsymbol{x'}$.
    return NO BUY
  end if
  Remove x'_{t-n} from list of most recent prices $\boldsymbol{x'}$.
  Add x_t^i to list of most recent prices $\boldsymbol{x'}$.
  Estimate the mean μ, and variance σ of last n returns.
  if $x_t^i < \mu - \tau * \sigma$ **then**
    return BUY
  else
    return NO BUY
  end if
end for

Final Result (2017)



All stocks



Without AAPL, UNH

Final Result (1998)

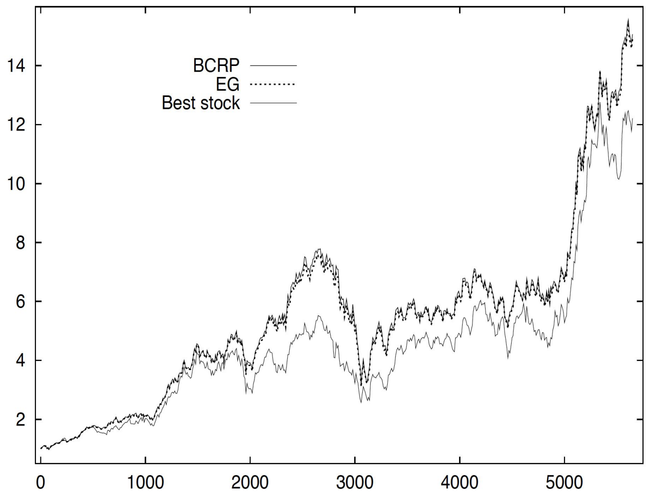


Figure 2: Comparison of wealths achieved by the best constant-rebalanced portfolio, the EG(η)-update and the best stock for a portfolio consisting of IBM and Coca Cola.

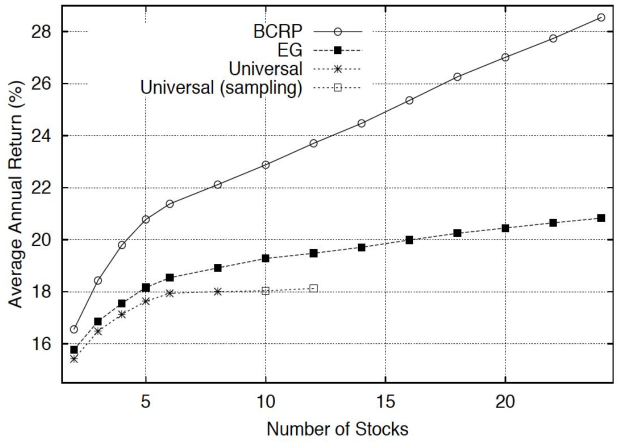


Figure 6: The average annual percent yield of BCRP, the universal portfolio algorithm and the EG(η)-update for random subsets of stocks from size 2 to 24.

EG: Then and Now

- EG still “works” — just not as well as it did for the original authors in 1998.
- Possible reasons:
  - Economic/historical reasons
  - Structural changes in financial markets
  - Efficient Markets Hypothesis — EG getting “priced in” after 1998?

	APY (%)					
	1998 (N=24)	*2017 (N=28)*	*2017 (N=26)**	*1998 (N=10)*	*2017 (N=10)*	*2017 (N=10)**
BCRP	~28.5%	26.8%	15.4%	~23%	26.8%	12.1%
EG	~21%	10.9%	8.6%	~19.5%	13.5%	7.7%

Additional Considerations

- Portfolio allocation in real life
  - Can only transact in integer number of shares
  - Transaction costs (fees, slippage, etc.)
- Incremental rebalancing
  - Reallocation step in Mean Reversion/Momentum strategies => large portfolio updates
- Potential survivorship bias in selecting current Dow Jones stocks
  - Or "reverse" survivorship bias (e.g. AAPL)
- More data
  - Evaluate algorithms on random subsets from larger set of stocks
  - What about other markets?
- Effect of side information on EG

Concluding Remarks

- We show that an online portfolio selection algorithm that maximizes a regularized cumulative reward function is able to achieve higher overall returns than a naive benchmark on modern-day, large-cap US equity data.
- We compare our results with experiments from the original EG paper and show a decrease in performance using post-2000 large-cap equity data, and observe the impact of outliers and a "winner-take-all" phenomenon.
- We show the potential for extending EG using "sophisticated" experts such as moving average-based indicators.
- We set up a custom simulation environment for running online learning algorithms on financial market data.
- Potential for further work:
    - Theoretical analysis of different sophisticated expert strategies
    - Incorporating side information
    - Shifting regret bounds

Questions