Trading Without Regret*

Michael Kearns
University of Pennsylvania

Nomura Quantitative Investment Strategies Conference
June 9, 2016

"No-Regret" Learning

- Have a set of N "signals" or "predictors"
  - alphas, advisors, returns of funds, long/short instruments,…
- Each trading period, each predictor receives an *arbitrary* payoff
  - *no* stochastic assumptions; could be generated by all-knowing adversary
  - predictors could have side information, expertise, specialization, omniscience, etc.
  - will assume boundedness of payoffs
- Algorithm maintains dynamic weighting/portfolio over predictors
  - receives weighted payoff each period
- Goal: over T periods, payoff close to the *best predictor in hindsight*
  - for *any* sequence of predictor payoffs
  - "no regret": (best payoff – algo payoff) grows *sublinearly* in T (→0 per period)
  - *not* competing with optimal (switching) policy
- Contrast with boosting criterion
  - trying to *track* best, not to *beat* best
  - "needle in a haystack" vs. "wisdom of crowds"
- Obvious appeal in financial settings

Multiplicative Weights algorithm

Initialization: Fix an $\eta \leq \frac{1}{2}$. With each decision i, associate the weight $w_i^{(1)} := 1$.

For $t = 1, 2, \ldots, T$:

1. Choose decision i with probability proportional to its weight $w_i^{(t)}$. I. e., use the distribution over decisions $\mathbf{p}^{(t)} = \{w_1^{(t)}/\Phi^{(t)}, \ldots, w_n^{(t)}/\Phi^{(t)}\}$ where $\Phi^{(t)} = \sum_i w_i^{(t)}$.
2. Observe the costs of the decisions $\mathbf{m}^{(t)}$.
3. Penalize the costly decisions by updating their weights as follows: for every decision i, set

$$w_i^{(t+1)} = w_i^{(t)}(1 - \eta m_i^{(t)}) \tag{2.1}$$

[Arora, Hazan, Kale]

Theorem 4.6 *The* Polynomial Weights (PW) *algorithm, using $\eta \leq 1/2$, for any $[0,1]$-valued loss sequence and for any k has,*

$$L_{\mathrm{PW}}^T \leq L_k^T + \eta Q_k^T + \frac{\ln(N)}{\eta},$$

where $Q_k^T = \sum_{t=1}^T (\ell_k^t)^2$. Setting $\eta = \min\{\sqrt{(\ln N)/T}, 1/2\}$ and noting that $Q_k^T \leq T$, we have $L_{\mathrm{PW}}^T \leq L_{min}^T + 2\sqrt{T \ln N}$.†

[Blum & Mansour]

Remarks

- No-regret algos and analyses have long and rich history
  - 1950s: Blackwell approachability
  - modern connections to Black-Scholes (non-stochastic derivation)
- Strong connections to game theory
  - minimax theorem and linear programming
  - convergence to Nash and correlated equilibrium
- Demystification #1: "Follow the Leader" has regret ~ # of lead changes
- Demystification #2: log(N) regret term means cannot try "everything"
  - e.g. can't predict sequence of T coin flips by adding all 2^T possible predictors
- Under stochastic assumptions, often recover (near) optimal solutions

Mathematical Finance, Vol. 1, No. 1 (January 1991), 1–29

UNIVERSAL PORTFOLIOS

THOMAS M. COVER[1]

Departments of Statistics and Electrical Engineering, Stanford University, Stanford, CA

We exhibit an algorithm for portfolio selection that asymptotically outperforms the best stock in the market. Let $\mathbf{x}_i = (x_{i1}, x_{i2}, \ldots, x_{im})^t$ denote the performance of the stock market on day i, where x_{ij} is the factor by which the jth stock increases on day i. Let $\mathbf{b}_i = (b_{i1}, b_{i2}, \ldots, b_{im})^t$, $b_{ij} \geq 0$, $\Sigma_j b_{ij} = 1$, denote the proportion b_{ij} of wealth invested in the jth stock on day i. Then $S_n = \Pi_{i=1}^n \mathbf{b}_i^t \mathbf{x}_i$ is the factor by which wealth is increased in n trading days. Consider as a goal the wealth $S_n^* = \max_{\mathbf{b}} \Pi_{i=1}^n \mathbf{b}^t \mathbf{x}_i$ that can be achieved by the best constant rebalanced portfolio chosen *after* the stock outcomes are revealed. It can be shown that S_n^* exceeds the best stock, the Dow Jones average, and the value line index at time n. In fact, S_n^* usually exceeds these quantities by an exponential factor. Let $\mathbf{x}_1, \mathbf{x}_2, \ldots,$ be an arbitrary sequence of market vectors. It will be shown that the nonanticipating sequence of portfolios $\hat{\mathbf{b}}_k = \int \mathbf{b} \Pi_{i=1}^{k-1} \mathbf{b}^t \mathbf{x}_i \, d\mathbf{b} / \int \Pi_{i=1}^{k-1} \mathbf{b}^t \mathbf{x}_i \, d\mathbf{b}$ yields wealth $\hat{S}_n = \Pi_{k=1}^n \hat{\mathbf{b}}_k^t \mathbf{x}_k$ such that $(1/n) \ln(S_n^*/\hat{S}_n) \to 0$, for every bounded sequence $\mathbf{x}_1, \mathbf{x}_2, \ldots,$ and, under mild conditions, achieves

$$\hat{S}_n \sim \frac{S_n^*(m-1)!(2\pi/n)^{(m-1)/2}}{|J_n|^{1/2}},$$

where J_n is an $(m-1) \times (m-1)$ sensitivity matrix. Thus this portfolio strategy has the same exponential rate of growth as the apparently unachievable S_n^*.

KEYWORDS: portfolio selection, robust trading strategies, performance weighting, rebalancing

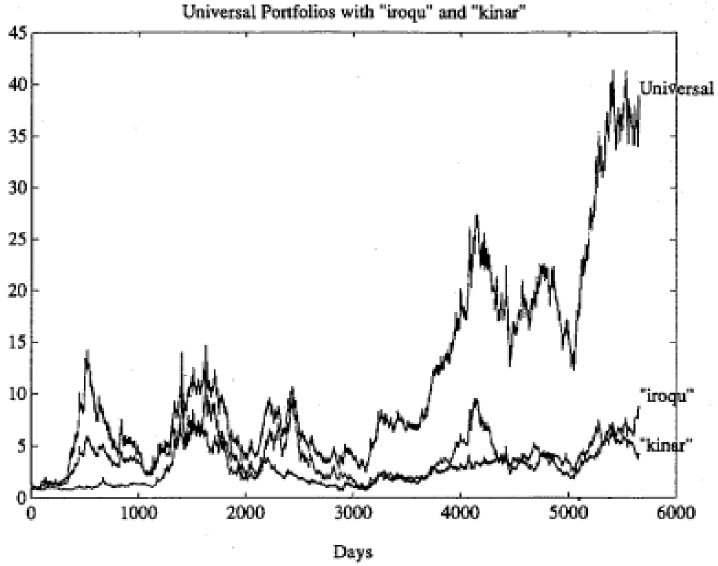


FIGURE 8.3. Performance of universal portfolio.

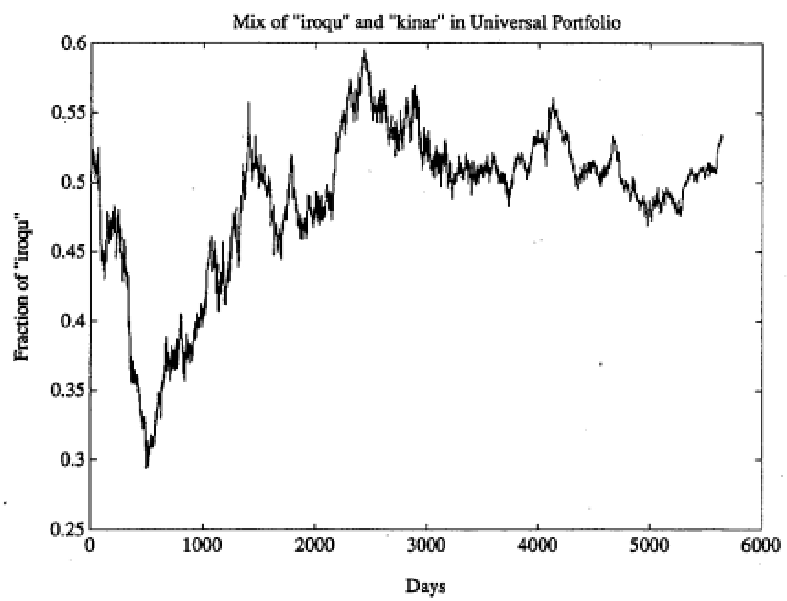


FIGURE 8.4. The portfolio $\hat{b}_k$.

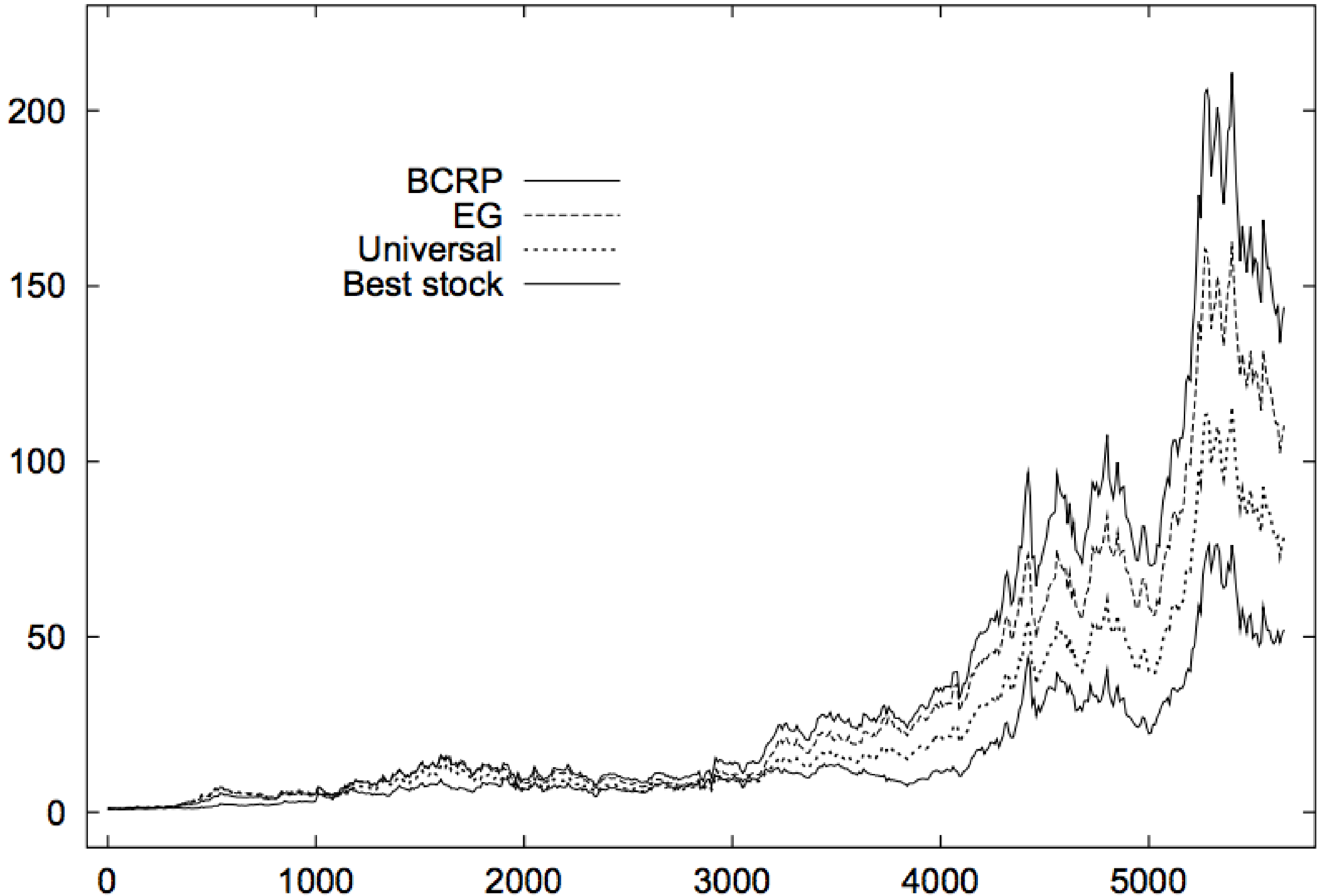
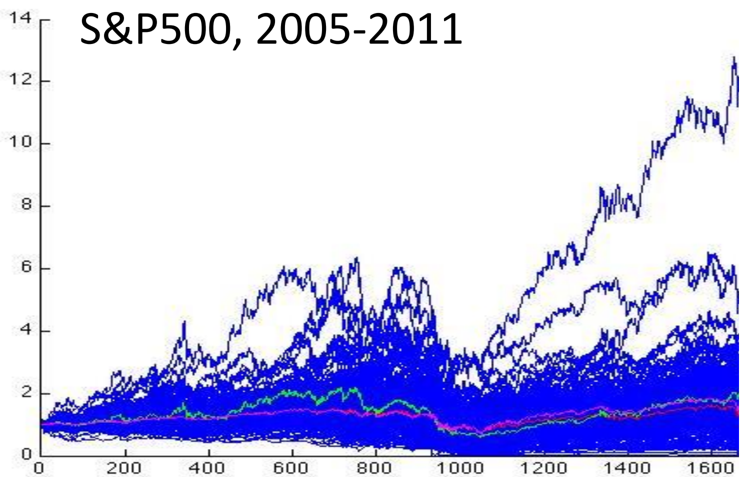


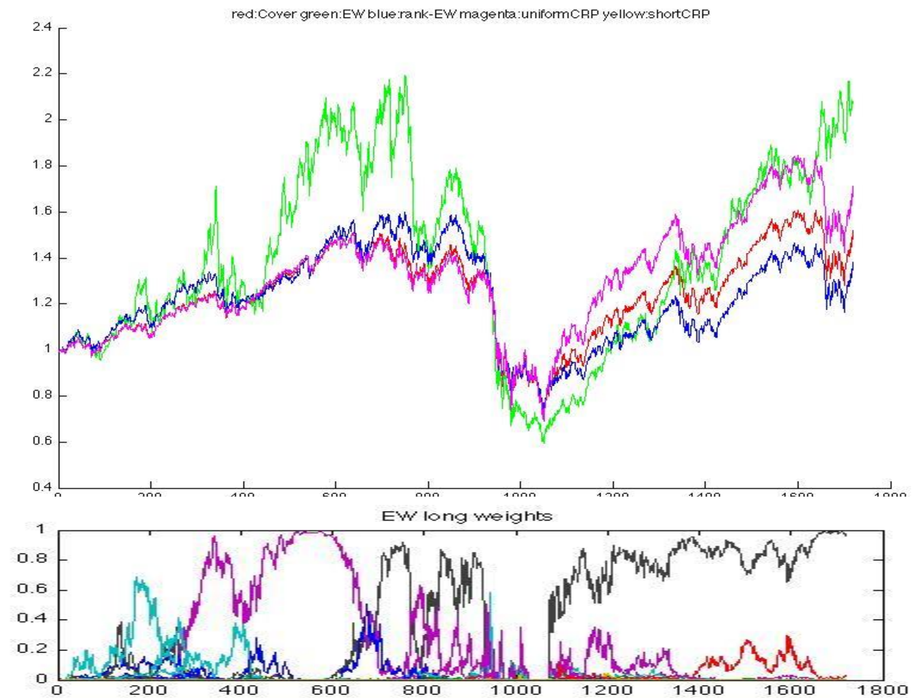
Figure 1: Comparison of wealths achieved by the best constant-rebalanced portfolio, the EG(η)-update, the universal portfolio algorithm and the best stock for a portfolio consisting of Commercial Metals and Kin Ark.

[Helmbold, Schapire, Singer, Warmuth]

Unfortunately…

- Large N and sideways markets create serious challenges
- Main issue: multiplicative updates lead to portfolio *concentration*
- Additive loss functions → lack of *risk* considerations
- Fiddling with learning rate doesn't help



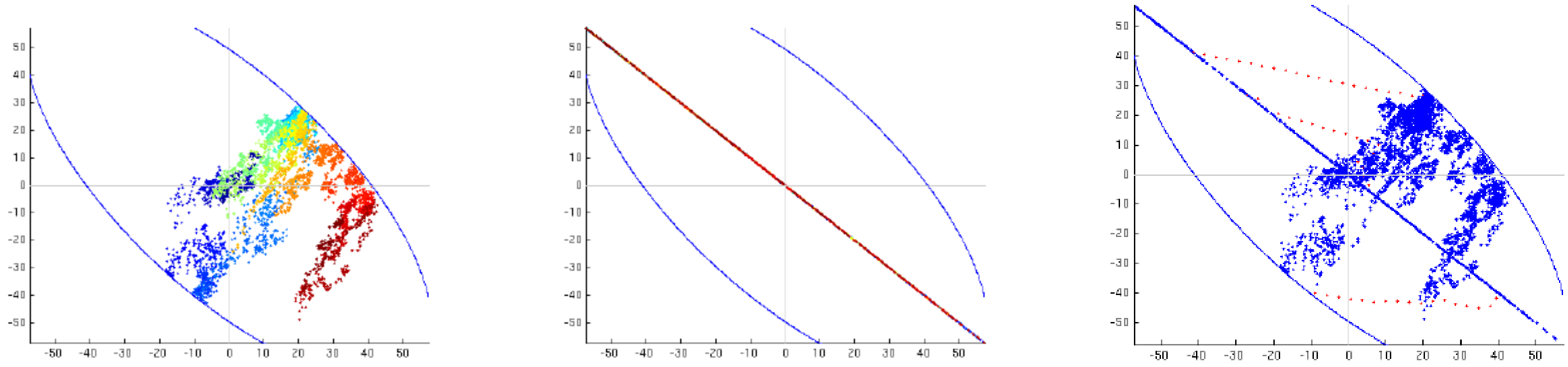


Even Worse…

- Could ask for no regret to best strategy in *risk-adjusted metrics:*
  - Sharpe Ratio: μ(returns)/σ(returns)
  - Mean-Variance: μ(returns) - σ(returns)
- Strong negative results:
  - No-regret *provably impossible*
  - Lower bounds on *competitive ratio* for any algorithm
- Intuition: Volatility introduces switching costs
- Alternative approach:
  - Measure risk by typical loss (e.g. one standard deviation)
  - *Internalize* risk within strategies

No-Regret Under Inventory Constraints

- Can't control Sharpe Ratio, but can limit allowed positions/portfolios
- Restrict to portfolios with daily standard deviation PNL at most $X historically
- Leads to elliptical constraint in portfolio space depending on correlations
- Only compete with strategies:
  - Obeying inventory constraints
  - Making only local moves (limit market impact)
- Combine no-regret with pursuit-evasion to recover (theoretical) guarantees



[Dworkin, K., Nevmyvaka]

Conclusions

- No-regret learning: rich history, powerful theory
- Deviations from additive loss (e.g. risk) present difficulties
- One workaround: endogenize risk
- Other uses: parameter optimization

Contact: mkearns@cis.upenn.edu