Online Portfolio Selection with Group Sparsity

Puja Das, Nicholas Johnson, Arindam Banerjee

Department of Computer Science and Engineering

University of Minnesota

Minneapolis, MN

Section 1: What is the problem?

Problem: How to invest in a few top performing sectors while accounting for transaction costs in online portfolio selection.

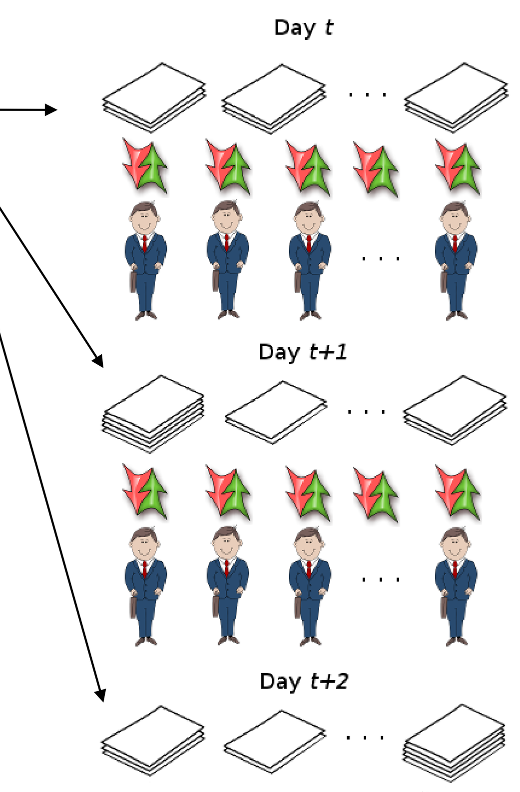
- Investors group stocks as sectors by the type of business.
- Not all sectors can yield profit.
- Sectors react differently during economic conditions.

Goal: Learn the underlying structure within stocks to identify top performing sectors.

Sector	Example Companies
Consumer Discretionary	Nike Inc., Target Corp.
Consumer Staples	Costco Co., Beam Inc.
Energy	Chevron Corp., Noble Corp.
Financials	Equifax Inc., AFLAC Inc.
Health Care	Cerner, Pfizer Inc.
Industrials	Raytheon Co., 3M Co.
Information Tech	Apple Inc., Dell Inc.
Materials	Alcoa Inc., Ecolab Inc.
Utilities	AGL Resources, AES Corp.

Section 2: Online Portfolio Selection

- Choose portfolio based on past stock performance: $\mathbf{p}_t = \langle p_t(1), \dots, p_t(n) \rangle$
- Price relatives $x_t(i)$ are the multiplicative factor by which a stock price changes
  - Loss: $x_t(i) < 1$ implies a loss
  - Gain: $x_t(i) > 1$ implies a gain
  - $x_t(i) = 1$ implies the price remained unchanged
- Maximize log gain in wealth: $LS_T(\mathbf{p}_{1:T}, \mathbf{x}_{1:T}) = \sum_{t=1}^{T} \log(\mathbf{p}_t^T \mathbf{x}_t)$



Section 3: What is the solution?

Solution: Group sparsity inducing regularizer in an online framework.

- At time t *select* p_t such that the regret is sublinear in T: $R_T = \sum_{t=1}^{T} \psi_t(\mathbf{p}_t) - \min_{\mathbf{p}^* \in \mathcal{P}} \sum_{t=1}^{T} \psi_t(\mathbf{p}^*) \leq o(T)$
  where $\psi_t(\mathbf{p}) = f_t(\mathbf{p}) + \lambda_1 \Omega(\mathbf{p}) + \lambda_2 \|\mathbf{p} - \mathbf{p}_{t-1}\|_1$
  - $f_t(.)$ -- convex loss function at time t
  - $\Omega(\cdot)$ -- groupwise *L2* norm for group sparsity $\Omega(\mathbf{p}) = \sum_{g=1}^{\mathcal{G}} w_g \|\mathbf{p}_{|g}\| \quad \forall g \in \mathcal{G}, g \subseteq \{1, \cdots, n\}$
  - $\|.\|_1$ -- *L1* norm inducing lazy updates
- Composite objective consisting of smooth and non-smooth terms.
- Can pose online portfolio selection as special case with $f_t(\mathbf{p}) = -\log(\mathbf{p}^T \mathbf{x}_t)$

Section 4: Online Lazy Updates with Group Sparsity (OLU-GS)

OLU-GS objective function:

$$\mathbf{p}_{t+1} = \underset{\mathbf{p} \in \Delta_n}{\operatorname{argmin}} \underbrace{\langle \nabla f_t(\mathbf{p}_t), \mathbf{p} \rangle + \lambda_1 \Omega(\mathbf{p}) + \lambda_2 \|\mathbf{p} - \mathbf{p}_t\|_1 + \frac{1}{2\eta} \|\mathbf{p} - \mathbf{p}_t\|_2^2}_{\text{Composite objective with lazy updates}}$$

- **ADMM Updates**

$$\mathbf{p}_{t+1}^{(k+1)} = \underset{\mathbf{p} \in \Delta_n}{\operatorname{argmin}} \langle \nabla f_t(\mathbf{p}_t), \mathbf{p} \rangle + \frac{1}{2\eta} \|\mathbf{p} - \mathbf{p}_t\|_2^2 + \frac{\beta}{2} \|\mathbf{p} - \mathbf{y}^{(k)} + \mathbf{w}^{(k)}\|_2^2 + \frac{\beta}{2} \|\mathbf{p} - \mathbf{p}_t - \mathbf{z}^{(k)} + \mathbf{v}^{(k)}\|_2^2$$

$$\mathbf{y}^{(k+1)} = \underset{\mathbf{y}}{\operatorname{argmin}} \ \lambda_1 \Omega(\mathbf{y}) + \frac{\beta}{2} \|\mathbf{p}_{t+1}^{(k+1)} - \mathbf{y} + \mathbf{w}^{(k)}\|_2^2 \rightarrow \forall g \in \mathcal{G}, \mathbf{y}_{|g} = \begin{cases} 0 & \text{if } \|\mathbf{q}_{|g}\|_2 \leq \tilde{\lambda} \\ \frac{\|\mathbf{q}_{|g}\|_2 - \tilde{\lambda}}{\|\mathbf{q}_{|g}\|_2} \mathbf{q}_{|g} & \text{otherwise} \end{cases}$$ **(Soft thresholding on groups of variables)**

$$\mathbf{z}^{(k+1)} = \underset{\mathbf{z}}{\operatorname{argmin}} \ \lambda_2 \|\mathbf{z}\|_1 + \frac{\beta}{2} \|\mathbf{p}_{t+1}^{(k+1)} - \mathbf{p}_t - \mathbf{z} + \mathbf{v}^{(k)}\|_2^2$$ **(Soft thresholding)**

$$\mathbf{w}^{(k+1)} = \mathbf{w}^{(k)} + (\mathbf{p}_{t+1}^{(k+1)} - \mathbf{y}^{(k+1)})$$ **(Closed form)**

$$\mathbf{v}^{(k+1)} = \mathbf{v}^{(k)} + (\mathbf{p}_{t+1}^{(k+1)} - \mathbf{p}_t - \mathbf{z}^{(k+1)})$$ **(Closed form)**

Analysis

- f_t is general convex: $R_T \leq O(\sqrt{T})$
- f_t is strongly convex: $R_T \leq O(\log(T))$

Section 5: Experiments and Results

- **Datasets: S&P500 (258 stocks,1991-2012): 9 sectors**
  NYSE (36 stocks,1962-1984): 8 sectors

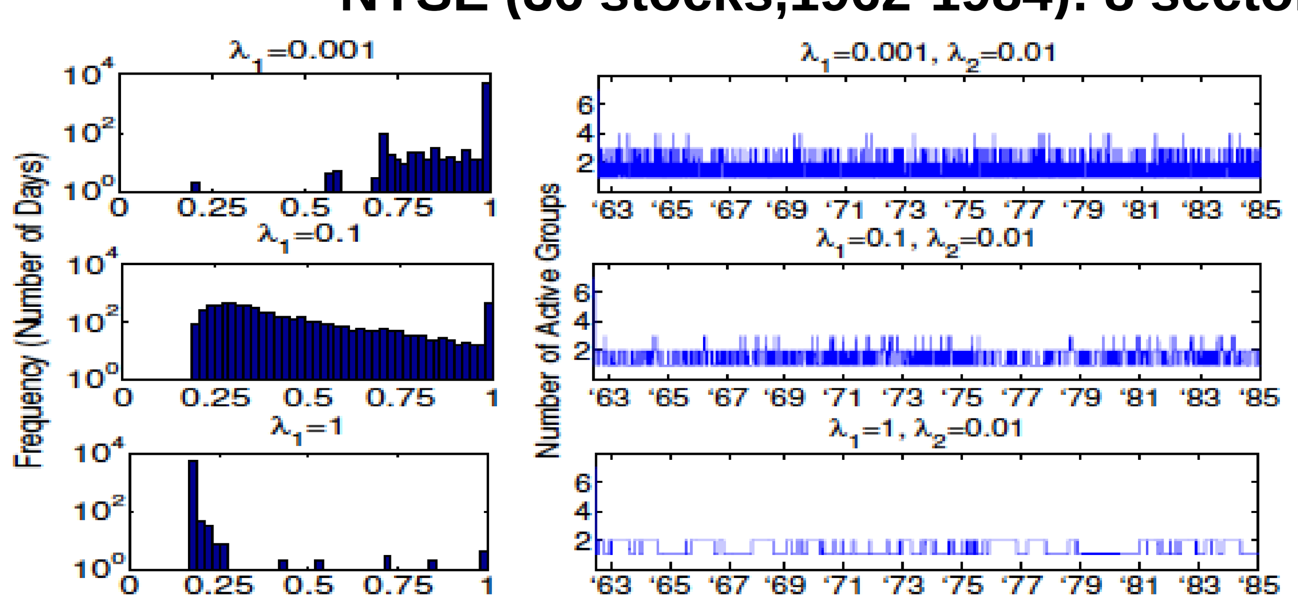


Figure 1: As λ_1 increases, the number of days with high group lasso value and active groups decrease.

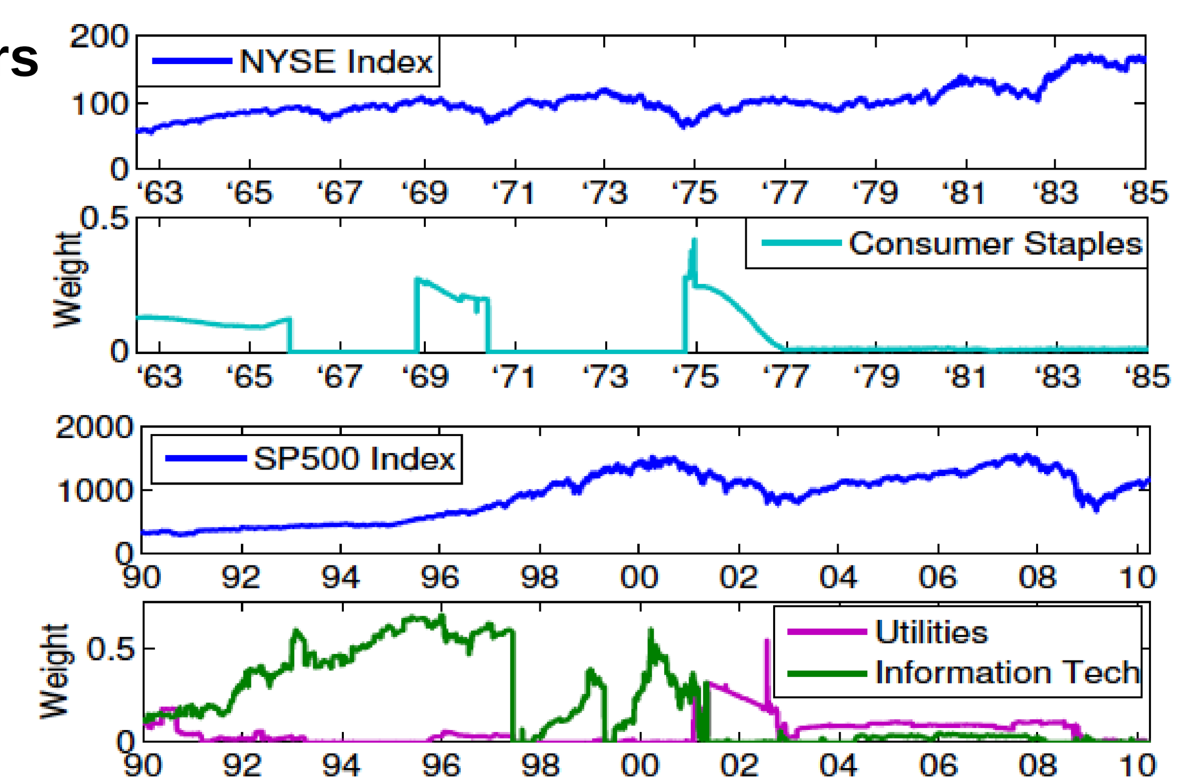


Figure 2: Cyclic and Non-cyclic sectors during bear and bull market.

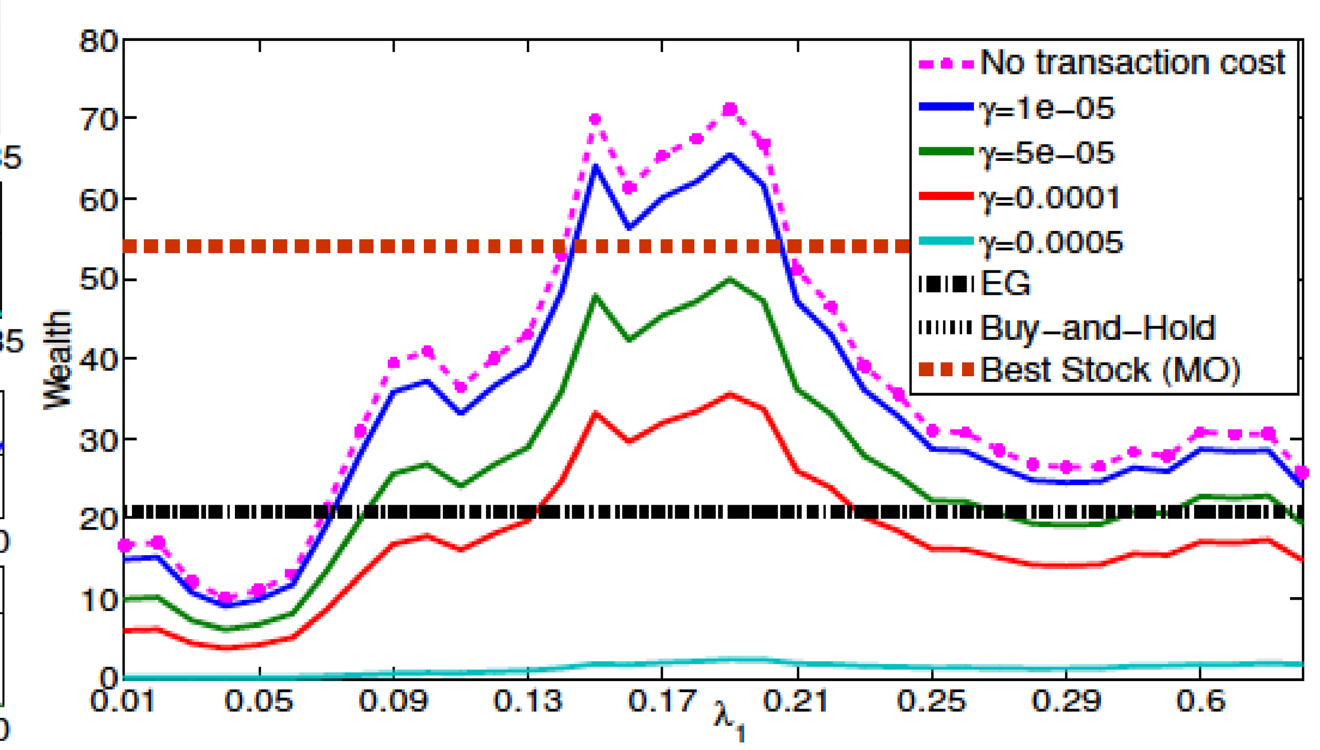


Figure 3: Transaction cost-adjusted wealth.

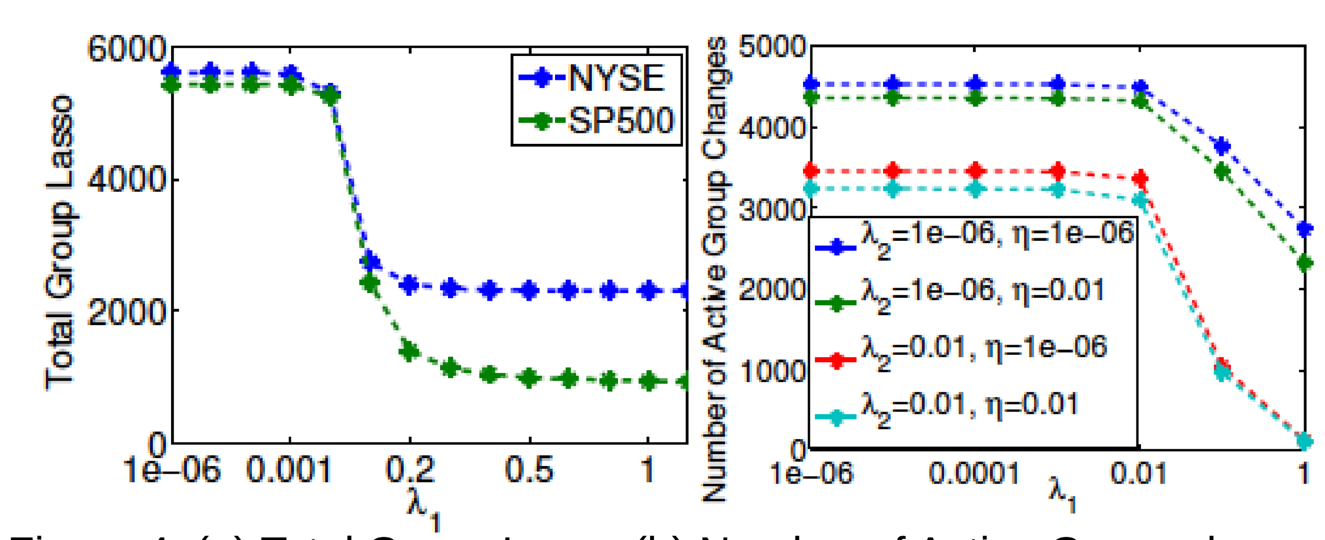


Figure 4: (a) Total Group Lasso (b) Number of Active Group changes.

Acknowledgements: The research was supported by NSF CAREER award IIS-0953274, NSF grants IIS-0916750, and IIS-1029711, and NASA grant NNX12AQ39A.