Online Resource Allocation with Structured Diversification

Nicholas Johnson

University of Minnesota

Joint work with Arindam Banerjee.

SIAM International Conference on Data Mining (SDM) 2015.

Resource Allocation

- Problem: Given n objects the goal is to find a strategy $\mathbf{p}$ which determines how to split up a resource over the n objects.
  - Resources: people, CPUs, money, and products.
  - Objects: software teams, server requests, assets, and warehouses.
- Examples: Job scheduling for compute servers, advertisement and recommendation serving, portfolio selection.
- Such problems often need to be solved online.

Resource Allocation

- Example: Online Portfolio Selection
- Objects are n different assets (stocks, options, bonds) and $\mathbf{p} \in \Delta_n = \{\mathbf{p} \,|\, p_i \geq 0 \sum_i p_i = 1\}$ is an investment strategy.
- Goal: Adaptively update a portfolio over a set of assets to maximize returns and minimize risk.

Diversification

- Reduce risk via diversification, i.e., avoid investing all of one's money in a few stocks.
- Naive strategy: $p_i \leq \kappa \; \forall i$ and $\kappa \in [0, 1]$.
- Example: Total: \$20, invest \$10 in Google stock and \$10 in Apple stock.
    - May not accomplish the goals of diversification because Google and Apple are structurally related.
- Must consider groups of assets that are structurally related and diversify across them.

$L_{(\infty,1)}$ Group Norm

- Let $\mathcal{G} = \{g_1, \ldots, g_m\}$ where $g_i = \{i_1, \ldots, i_{n_i}\}$ is a set of assets in g_i.
  - $|g_i| = n_i$, and $g_i \cap g_j \neq \emptyset$.
- Let $\mathbf{x} \in \mathbb{R}^n$, then $\mathbf{x}_{g_i}$ is a sub-vector with values equal to $\mathbf{x}$ for indices in g_i and 0 otherwise.
- Our contribution (1): We propose the $L_{(\infty,1)}$ group norm which plays a key role in structured diversity

$$L_{(\infty,1)} := \left\| \left[\|\mathbf{x}_{g_1}\|_1 \ldots \|\mathbf{x}_{g_m}\|_1 \right]^\top \right\|_\infty . \quad (1)$$

$L_{(\infty,1)}$ Group Norm

0.08	0.11	0.06	0.01	0.02	0.12	0.04	0.07	0.21	0.28

$L_{(\infty,1)}$ Group Norm

0.08	0.11	0.06	0.01	0.02	0.12	0.04	0.07	0.21	0.28

$\underbrace{\hspace{4cm}}_{\mathbf{x}_{g_1}}$

$L_{(\infty,1)}$ Group Norm

0.08	0.11	0.06	0.01	0.02	0.12	0.04	0.07	0.21	0.28

$\underbrace{\hphantom{0.08 \quad 0.11 \quad 0.06 \quad 0.01}}_{\mathbf{x}_{g_1}}$

0.08	0.11	0.06	0.01

$L_{(\infty,1)}$ Group Norm

0.08	0.11	0.06	0.01	0.02	0.12	0.04	0.07	0.21	0.28

$\mathbf{x}_{g_1}$

0.08	0.11	0.06	0.01

$\mathbf{x}_{g_2}$

$L_{(\infty,1)}$ Group Norm

0.08	0.11	0.06	0.01	0.02	0.12	0.04	0.07	0.21	0.28

$\mathbf{x}_{g_1}$

0.08	0.11	0.06	0.01

$\mathbf{x}_{g_2}$

0.06	0.01	0.02	0.12	0.04	0.07

$L_{(\infty,1)}$ Group Norm

0.08	0.11	0.06	0.01	0.02	0.12	0.04	0.07	0.21	0.28

$\mathbf{x}_{g_1}$

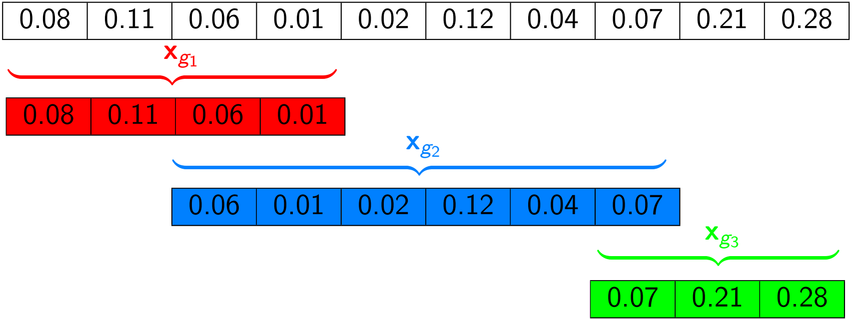
0.08	0.11	0.06	0.01

$\mathbf{x}_{g_2}$

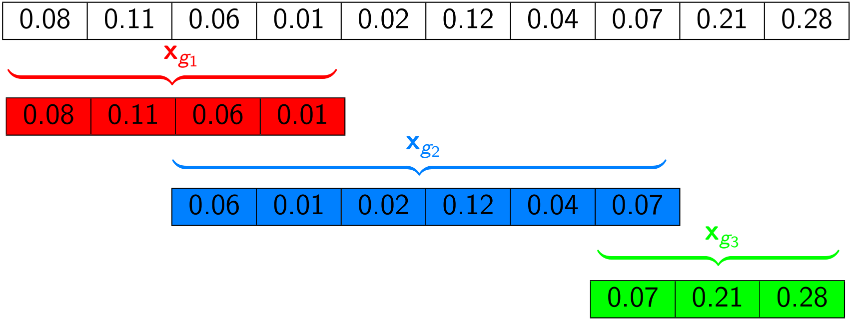
0.06	0.01	0.02	0.12	0.04	0.07

$\mathbf{x}_{g_3}$

$L_{(\infty,1)}$ Group Norm



$L_{(\infty,1)}$ Group Norm



- $\|\mathbf{x}_{g_1}\|_1 = 0.26$, $\|\mathbf{x}_{g_2}\|_1 = 0.32$, $\|\mathbf{x}_{g_3}\|_1 = 0.56$

$L_{(\infty,1)}$ Group Norm

0.08	0.11	0.06	0.01	0.02	0.12	0.04	0.07	0.21	0.28

$\mathbf{x}_{g_1}$

0.08	0.11	0.06	0.01

$\mathbf{x}_{g_2}$

0.06	0.01	0.02	0.12	0.04	0.07

$\mathbf{x}_{g_3}$

0.07	0.21	0.28

- $\|\mathbf{x}_{g_1}\|_1 = 0.26, \ \|\mathbf{x}_{g_2}\|_1 = 0.32, \ \|\mathbf{x}_{g_3}\|_1 = 0.56$
- $\left\| [0.26 \ 0.32 \ 0.56]^\top \right\|_\infty = \mathbf{0.56}$

Structured Diversity with $L_{(\infty,1)}$ Group Norm

- Use $L_{(\infty,1)}$ as constraint: $\|\mathbf{p}\|^{\mathcal{G}}_{(\infty,1)} \leq \kappa$.
- Constraint allows each group g_i to increase their L_1 norm to κ without changing the norm ball (penalty).
- We consider problems of the form

$$\min_{\mathbf{p}\in\Delta_n} \quad f(\mathbf{p}) \quad \text{s.t. } \|\mathbf{p}\|^{\mathcal{G}}_{(\infty,1)} \leq \kappa \tag{2}$$

where groups are assumed to be known.

Online Resource Allocation

- We pose the online resource allocation problem in the setting of Online Convex Optimization (OCO):
  - For $t = 1, \ldots, T$
    - Player chooses a solution $\mathbf{p}_t$ from convex set $\mathcal{X}$.
    - Nature chooses a convex loss function $\ell_t(\cdot)$.
    - The loss function is revealed and the player suffers a loss of $\ell_t(\mathbf{p}_t)$.

Online Resource Allocation with Structured Diversification (ORASD)

- Two requirements on $\mathbf{p}_t$:
  1. Keep $\mathbf{p}_t$ close to $\mathbf{p}_{t-1}$ using distance function Ω (Bregman Divergence).
  2. $\|\mathbf{p}_t\|_{(\infty,1)}^{\mathcal{G}} \leq \kappa$ so structural diversity is maintained over time.

Online Resource Allocation with Structured Diversification (ORASD)

- Two requirements on $\mathbf{p}_t$:
  1. Keep $\mathbf{p}_t$ close to $\mathbf{p}_{t-1}$ using distance function Ω (Bregman Divergence).
  2. $\|\mathbf{p}_t\|^{\mathcal{G}}_{(\infty,1)} \leq \kappa$ so structural diversity is maintained over time.

- The problem is then

$$\min_{\mathbf{p}\in\Delta_n} \eta\,\ell_t(\mathbf{p}) + \Omega(\mathbf{p}, \mathbf{p}_{t-1}) \quad \text{s.t. } \|\mathbf{p}\|^{\mathcal{G}}_{(\infty,1)} \leq \kappa \tag{3}$$

where $\eta \geq 0$.

Online Resource Allocation with Structured Diversification (ORASD)

- We solve a linearized version using a first-order Taylor expansion of f_t at $\mathbf{p}_t$ along with a proximal term.
- Our contribution (2): We present an online resource allocation framework for structured diversification (ORASD)

$$\mathbf{p}_{t+1} := \operatorname*{argmin}_{\substack{\mathbf{p}\in\Delta_n \\ \|\mathbf{p}\|^{\mathcal{G}}_{(\infty,1)}\leq\kappa}} \eta\langle \mathbf{p}, \nabla\ell_t(\mathbf{p}_t)\rangle + \frac{1}{2}\|\mathbf{p}-\mathbf{p}_t\|_2^2 + \lambda\Omega(\mathbf{p},\mathbf{p}_t) \tag{4}$$

where $\lambda \geq 0$.

Online Portfolio Selection

- Online Portfolio Selection is a special case of (4).
- Consider a stock market consisting of n stocks $\{s(1), \ldots, s(n)\}$ over a span of T days.
- Price relatives: $x_t(i) = \frac{\text{closing price}}{\text{opening price}}$, $\mathbf{x}_t = [x_t(1) \ldots x_t(n)]^\top$.
- Portfolio: $\mathbf{p}_t = [p_t(1) \ldots p_t(n)]^\top \in \Delta_n$ prescribes investing $p_t(i)$ fraction of the current wealth in stock $s_t(i)$.
- Multiplicative gain in wealth for day t: $\mathbf{p}_t^\top \mathbf{x}_t$.
- Logarithmic gain in wealth over T days: $\sum_{t=1}^{T} \log(\mathbf{p}_t^\top \mathbf{x}_t)$.

Online Portfolio Selection with ORASD

- $\ell_t(\mathbf{p}) = -\log(\mathbf{p}^\top \mathbf{x}_t)$.
- $\Omega = \|\mathbf{p} - \mathbf{p}_t\|_1$.
- The online portfolio selection with structured diversification problem is

$$\min_{\substack{\mathbf{p}\in\Delta_n \\ \|\mathbf{p}\|^{\mathcal{G}}_{(\infty,1)}\leq\kappa}} \quad \eta\langle\mathbf{p}, \nabla\ell_t(\mathbf{p}_t)\rangle + \frac{1}{2}\|\mathbf{p}-\mathbf{p}_t\|_2^2 + \lambda\|\mathbf{p}-\mathbf{p}_t\|_1 \ . \tag{5}$$

- The L_1 penalty term measures the fraction of wealth traded.
- λ controls the amount that can be traded every day.
- The $L_{(\infty,1)}$ constraint induces diversification where κ is the level of diversification desired.

Online Portfolio Selection with ORASD

- We use lifting in solving (5) by treating each $\mathbf{p}_{g_i}$ as a separate variable and letting $\|\mathbf{p}\|_{(\infty,1)}^{\mathcal{G}} \leq \kappa \equiv \|\mathbf{p}_{g_i}\|_1 \leq \kappa \; \forall i$.
- We determine the final $\mathbf{p}$ to use by requiring the overlapping indices to be the same via a consensus constraint.
- Consensus constraint: $\mathbf{p}|_{g_i} = \mathbf{p}_{g_i} \; \forall i$.
  - $\mathbf{p}|_{g_i}$ is equal to $\mathbf{p}$ for indices in g_i.
  - $\mathbf{p}_{g_i}$ is separate variable we compute.
  - These must be equal at convergence.

Online Portfolio Selection with ORASD

- $k = 1$: $\underbrace{\begin{bmatrix} 13 \\ 2 \\ 5 \\ 6 \end{bmatrix}}_{\mathbf{p}}, \underbrace{\begin{bmatrix} 13 \\ 0 \\ 5 \\ 6 \end{bmatrix}}_{\mathbf{p}|_{g_1}} = \underbrace{\begin{bmatrix} 7 \\ 0 \\ 13 \\ 40 \end{bmatrix}}_{\mathbf{p}_{g_1}}, \underbrace{\begin{bmatrix} 13 \\ 2 \\ 5 \\ 0 \end{bmatrix}}_{\mathbf{p}|_{g_2}} = \underbrace{\begin{bmatrix} 1 \\ 12 \\ 9 \\ 0 \end{bmatrix}}_{\mathbf{p}_{g_2}}.$

Online Portfolio Selection with ORASD

- $k = 1$: $\underbrace{\begin{bmatrix}13\\2\\5\\6\end{bmatrix}}_{\mathbf{p}}, \underbrace{\begin{bmatrix}13\\0\\5\\6\end{bmatrix}}_{\mathbf{p}|_{g_1}} = \underbrace{\begin{bmatrix}7\\0\\13\\40\end{bmatrix}}_{\mathbf{p}_{g_1}}, \underbrace{\begin{bmatrix}13\\2\\5\\0\end{bmatrix}}_{\mathbf{p}|_{g_2}} = \underbrace{\begin{bmatrix}1\\12\\9\\0\end{bmatrix}}_{\mathbf{p}_{g_2}}.$
- $k = 2$: $\underbrace{\begin{bmatrix}10\\12\\23\\4\end{bmatrix}}_{\mathbf{p}}, \underbrace{\begin{bmatrix}10\\0\\23\\4\end{bmatrix}}_{\mathbf{p}|_{g_1}} = \underbrace{\begin{bmatrix}10\\0\\21\\14\end{bmatrix}}_{\mathbf{p}_{g_1}}, \underbrace{\begin{bmatrix}10\\12\\23\\0\end{bmatrix}}_{\mathbf{p}|_{g_2}} = \underbrace{\begin{bmatrix}9\\12\\27\\0\end{bmatrix}}_{\mathbf{p}_{g_2}}.$

Online Portfolio Selection with ORASD

- $k = 1$: $\underbrace{\begin{bmatrix}13\\2\\5\\6\end{bmatrix}}_{\mathbf{p}}, \underbrace{\begin{bmatrix}13\\0\\5\\6\end{bmatrix}}_{\mathbf{p}|_{g_1}} = \underbrace{\begin{bmatrix}7\\0\\13\\40\end{bmatrix}}_{\mathbf{p}_{g_1}}, \underbrace{\begin{bmatrix}13\\2\\5\\0\end{bmatrix}}_{\mathbf{p}|_{g_2}} = \underbrace{\begin{bmatrix}1\\12\\9\\0\end{bmatrix}}_{\mathbf{p}_{g_2}}.$
- $k = 2$: $\underbrace{\begin{bmatrix}10\\12\\23\\4\end{bmatrix}}_{\mathbf{p}}, \underbrace{\begin{bmatrix}10\\0\\23\\4\end{bmatrix}}_{\mathbf{p}|_{g_1}} = \underbrace{\begin{bmatrix}10\\0\\21\\14\end{bmatrix}}_{\mathbf{p}_{g_1}}, \underbrace{\begin{bmatrix}10\\12\\23\\0\end{bmatrix}}_{\mathbf{p}|_{g_2}} = \underbrace{\begin{bmatrix}9\\12\\27\\0\end{bmatrix}}_{\mathbf{p}_{g_2}}.$

$$\vdots$$

Online Portfolio Selection with ORASD

- $k=1$: $\underbrace{\begin{bmatrix}13\\2\\5\\6\end{bmatrix}}_{\mathbf{p}}, \underbrace{\begin{bmatrix}13\\0\\5\\6\end{bmatrix}}_{\mathbf{p}|_{g_1}} = \underbrace{\begin{bmatrix}7\\0\\13\\40\end{bmatrix}}_{\mathbf{p}_{g_1}}, \underbrace{\begin{bmatrix}13\\2\\5\\0\end{bmatrix}}_{\mathbf{p}|_{g_2}} = \underbrace{\begin{bmatrix}1\\12\\9\\0\end{bmatrix}}_{\mathbf{p}_{g_2}}.$
- $k=2$: $\underbrace{\begin{bmatrix}10\\12\\23\\4\end{bmatrix}}_{\mathbf{p}}, \underbrace{\begin{bmatrix}10\\0\\23\\4\end{bmatrix}}_{\mathbf{p}|_{g_1}} = \underbrace{\begin{bmatrix}10\\0\\21\\14\end{bmatrix}}_{\mathbf{p}_{g_1}}, \underbrace{\begin{bmatrix}10\\12\\23\\0\end{bmatrix}}_{\mathbf{p}|_{g_2}} = \underbrace{\begin{bmatrix}9\\12\\27\\0\end{bmatrix}}_{\mathbf{p}_{g_2}}.$

$\vdots$

- $k=K$: $\underbrace{\begin{bmatrix}10\\20\\30\\40\end{bmatrix}}_{\mathbf{p}}, \underbrace{\begin{bmatrix}10\\0\\30\\40\end{bmatrix}}_{\mathbf{p}|_{g_1}} = \underbrace{\begin{bmatrix}10\\0\\30\\40\end{bmatrix}}_{\mathbf{p}_{g_1}}, \underbrace{\begin{bmatrix}10\\20\\30\\0\end{bmatrix}}_{\mathbf{p}|_{g_2}} = \underbrace{\begin{bmatrix}10\\20\\30\\0\end{bmatrix}}_{\mathbf{p}_{g_2}}$ at convergence.

Online Portfolio Selection with ORASD

- Since our problem consists of smooth and non-smooth terms, we use ADMM to efficiently solve it by introducing auxiliary variable $\mathbf{z} = \mathbf{p} - \mathbf{p}_t$ and $h(\mathbf{p}_{g_i}) = \mathbb{1}_{(\|\mathbf{p}_{g_i}\|_1 \leq \kappa)}$.
- The ADMM formulation of the problem is

$$\min_{\substack{\mathbf{p}\in\Delta_n \\ \mathbf{p}|_{g_i}=\mathbf{p}_{g_i}\ \forall i \\ \mathbf{p}-\mathbf{p}_t=\mathbf{z}}} \eta\langle \mathbf{p}, \nabla\ell_t(\mathbf{p}_t)\rangle + \frac{1}{2}\|\mathbf{p}-\mathbf{p}_t\|_2^2 + \lambda\|\mathbf{z}\|_1 + \sum_{i=1}^{m} h(\mathbf{p}_{g_i}) \ . \tag{6}$$

Online Portfolio Selection with ORASD

- The augmented Lagrangian of (6) is

$$L(\mathbf{p}, \mathbf{p}_{g_1 \dots g_m}, \mathbf{z}, \mathbf{u}_{1 \dots m}, \mathbf{v}) = \eta \langle \mathbf{p}, \nabla \ell_t(\mathbf{p}_t) \rangle + \frac{1}{2} \|\mathbf{p} - \mathbf{p}_t\|_2^2 + \lambda \|\mathbf{z}\|_1 + \sum_{i=1}^{m} h(\mathbf{p}_{g_i}) + \frac{\beta}{2} \sum_{i=1}^{m} \|\mathbf{p}|_{g_i} - \mathbf{p}_{g_i} + \mathbf{u}_i\|_2^2 + \frac{\beta}{2} \|\mathbf{p} - \mathbf{p}_t - \mathbf{z} + \mathbf{v}\|_2^2$$

where **u** and **v** are scaled dual variables.

Online Portfolio Selection with ORASD

- Compute optimal variables of the augmented Lagrangian by iteratively minimizing primal variables and maximizing dual variables:
  - For $k = 1, \ldots$
    - $\mathbf{p}^{k+1} = \underset{\mathbf{p}}{\operatorname{argmin}}\ L(\mathbf{p}, \mathbf{p}^k_{g_1 \ldots g_m}, \mathbf{z}^k, \mathbf{u}^k_{1 \ldots m}, \mathbf{v}^k)$
    - $\mathbf{p}^{k+1}_{g_1 \ldots g_m} = \underset{\mathbf{p}_{g_i}}{\operatorname{argmin}}\ L(\mathbf{p}^{k+1}, \mathbf{p}_{g_1 \ldots g_m}, \mathbf{z}^k, \mathbf{u}^k_{1 \ldots m}, \mathbf{v}^k)$
    - $\mathbf{z}^{k+1} = \underset{\mathbf{z}}{\operatorname{argmin}}\ L(\mathbf{p}^{k+1}, \mathbf{p}^{k+1}_{g_1 \ldots g_m}, \mathbf{z}, \mathbf{u}^k_{1 \ldots m}, \mathbf{v}^k)$
    - $\mathbf{u}^{k+1}_{1 \ldots m} = \underset{\mathbf{u}_i}{\operatorname{argmax}}\ L(\mathbf{p}^{k+1}, \mathbf{p}^{k+1}_{g_1 \ldots g_m}, \mathbf{z}^{k+1}, \mathbf{u}_{1 \ldots m}, \mathbf{v}^k)$ Dual ascent
    - $\mathbf{v}^{k+1} = \underset{\mathbf{v}}{\operatorname{argmax}}\ L(\mathbf{p}^{k+1}, \mathbf{p}^{k+1}_{g_1 \ldots g_m}, \mathbf{z}^{k+1}, \mathbf{u}^{k+1}_{1 \ldots m}, \mathbf{v})$ Dual ascent

ORASD: $\mathbf{p}$-update

$$\mathbf{p}^{k+1} := \operatorname{argmin}_{\mathbf{p}\in\Delta_n} \; \eta\langle \mathbf{p}, \nabla \ell_t(\mathbf{p}_t)\rangle + \frac{1}{2}\|\mathbf{p}-\mathbf{p}_t\|_2^2 + \frac{\beta}{2}\sum_{i=1}^{m}\|\mathbf{p}|_{g_i} - \mathbf{p}_{g_i}^k + \mathbf{u}_i^k\|_2^2 + \frac{\beta}{2}\|\mathbf{p}-\mathbf{p}_t-\mathbf{z}^k+\mathbf{v}^k\|_2^2 \,.$$

- We solve for $\mathbf{p}$ be taking the gradient *w.r.t.* $\mathbf{p}$ and setting it to zero to get the closed form update of $\mathbf{p}$ as

$$\mathbf{p} = \prod_{\Delta_n}\Big(\hat{\mathbf{a}}\frac{\eta \mathbf{x}_t}{\mathbf{p}_t^\top \mathbf{x}_t} + \hat{\mathbf{a}}(1+\beta)\mathbf{p}_t + \hat{\mathbf{a}}\beta\big(\mathbf{S}_{g_1}(\mathbf{p}_{g_1}^k - \mathbf{u}_1^k) + \ldots + \mathbf{S}_{g_m}(\mathbf{p}_{g_m}^k - \mathbf{u}_m^k) + \mathbf{z}^k - \mathbf{v}^k\big)\Big)$$

  where $\hat{\mathbf{a}} = \left((1+\beta)I + \beta\left(\mathbf{S}_{g_1} + \cdots + \mathbf{S}_{g_m}\right)\right)^{-1}$ and $\prod_{\Delta_n}$ is a projection to the probability simplex.

- Projected gradient descent step.

ORASD: $\mathbf{p}_{g_i}$-update

$$\mathbf{p}_{g_i}^{k+1} := \operatorname{argmin}_{\mathbf{p}_{g_i}} \mathbb{1}_{(\|\mathbf{p}_{g_i}\| \leq \kappa)} + \frac{\beta}{2} \|\mathbf{p}^{k+1}|_{g_i} - \mathbf{p}_{g_i} + \mathbf{u}_i^k\|_2^2 \ \forall i \ .$$

- We solve each $\mathbf{p}_{g_i}$ in parallel by projecting to the L_1 ball of radius κ

$$\mathbf{p}_{g_i} = \prod_{\|\cdot\|_1 \leq \kappa} \left(\mathbf{S}_{g_i} \mathbf{p}^{k+1} + \mathbf{u}_i^k \right) \ .$$

ORASD: $\mathbf{z}$-update

$$\mathbf{z}^{k+1} := \operatorname{argmin}_{\mathbf{z}} \lambda \|\mathbf{z}\|_1 + \frac{\beta}{2} \|\mathbf{p}^{k+1} - \mathbf{p}_t - \mathbf{z} + \mathbf{v}^k\|_2^2 \ .$$

- We solve for $\mathbf{z}$ by using the soft-thresholding operator $S_\rho(a)$

$$\mathbf{z} = S_{\lambda/\beta}(\mathbf{p}^{k+1} - \mathbf{p}_t + \mathbf{v}^k) \ .$$

ORASD: $\mathbf{u}_i$ and $\mathbf{v}$ updates

$$\mathbf{u}_i^{k+1} := \mathbf{u}_i^k + \mathbf{p}^{k+1}|_{g_i} - \mathbf{p}_{g_i}^{k+1} \; \forall i \; .$$

$$\mathbf{v}^{k+1} := \mathbf{v}^k + \mathbf{p}^{k+1} - \mathbf{p}_t - \mathbf{z}^{k+1} \; .$$

- Both $\mathbf{u}_i \; \forall i$ and $\mathbf{v}$ are already in closed form. We can compute $\mathbf{u}_i$ in parallel for all i.

Experiments

Initial investment of $1, uniformly invested over the stocks.
Datasets:

- New York Stock Exchange (NYSE): 36 stocks, data from 1962 - 1984, $\sim$ 6000 days.
- Standard & Poor's 500 (S&P500): 263 stocks, data from 1990 - 2012, $\sim$ 5500 days.

Groups:

- Compute correlation graph C. Remove edges with weight $< \epsilon$. For each stock, create a group around it by including all directly connected stocks.

Risk and κ

- Covariance: Compute risk covariance of $\mathbf{p}_t$ *w.r.t.* a uniform portfolio $\mathbf{u}$ as $\alpha_{cov} = \mathbf{p}_t^\top \boldsymbol{\Sigma}_t \mathbf{p}_t / \mathbf{u}^\top \boldsymbol{\Sigma}_t \mathbf{u}$.

Risk and κ

- Covariance: Compute risk covariance of $\mathbf{p}_t$ *w.r.t.* a uniform portfolio $\mathbf{u}$ as $\alpha_{cov} = \mathbf{p}_t^\top \boldsymbol{\Sigma}_t \mathbf{p}_t / \mathbf{u}^\top \boldsymbol{\Sigma}_t \mathbf{u}$.
  - High α_{cov} implies high risk and low α_{cov} implies low risk.

Risk and κ

- Covariance: Compute risk covariance of $\mathbf{p}_t$ *w.r.t.* a uniform portfolio $\mathbf{u}$ as $\alpha_{cov}=\mathbf{p}_t^\top \mathbf{\Sigma}_t \mathbf{p}_t / \mathbf{u}^\top \mathbf{\Sigma}_t \mathbf{u}$.
  - High α_{cov} implies high risk and low α_{cov} implies low risk.
- Sharpe ratio: Measures compensation w.r.t. a portfolio for risk taken for both downwards and upwards volatility as $\alpha_{sharpe}=(R-R_b)/\sqrt{var(R-R_b)}$ where R is the return for the portfolio and R_b is the benchmark return.

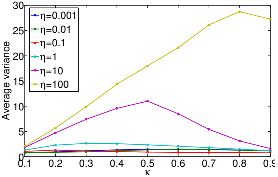
Risk and κ

- Covariance: Compute risk covariance of $\mathbf{p}_t$ *w.r.t.* a uniform portfolio $\mathbf{u}$ as $\alpha_{cov}{=}\mathbf{p}_t^\top \mathbf{\Sigma}_t \mathbf{p}_t / \mathbf{u}^\top \mathbf{\Sigma}_t \mathbf{u}$.
  - High α_{cov} implies high risk and low α_{cov} implies low risk.
- Sharpe ratio: Measures compensation w.r.t. a portfolio for risk taken for both downwards and upwards volatility as $\alpha_{sharpe}{=}(R{-}R_b)/\sqrt{var(R{-}R_b)}$ where R is the return for the portfolio and R_b is the benchmark return.
- Sortino ratio: Similar to the Sharpe ratio however only measures the downwards volatility computed as $\alpha_{sortino} = (R - R_b)/DR$ where DR is the standard deviation of negative returns (losses).

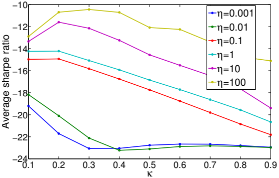
Risk and κ

- Covariance: Compute risk covariance of $\mathbf{p}_t$ *w.r.t.* a uniform portfolio $\mathbf{u}$ as $\alpha_{cov}=\mathbf{p}_t^\top \boldsymbol{\Sigma}_t \mathbf{p}_t / \mathbf{u}^\top \boldsymbol{\Sigma}_t \mathbf{u}$.
  - High α_{cov} implies high risk and low α_{cov} implies low risk.
- Sharpe ratio: Measures compensation w.r.t. a portfolio for risk taken for both downwards and upwards volatility as $\alpha_{sharpe}=(R-R_b)/\sqrt{var(R-R_b)}$ where R is the return for the portfolio and R_b is the benchmark return.
- Sortino ratio: Similar to the Sharpe ratio however only measures the downwards volatility computed as $\alpha_{sortino} = (R - R_b)/DR$ where DR is the standard deviation of negative returns (losses).
  - Low Sharpe and Sortino ratio implies high risk for given return.

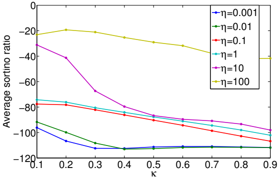
Risk and κ



Covariance



Sharpe ratio



Sortino ratio

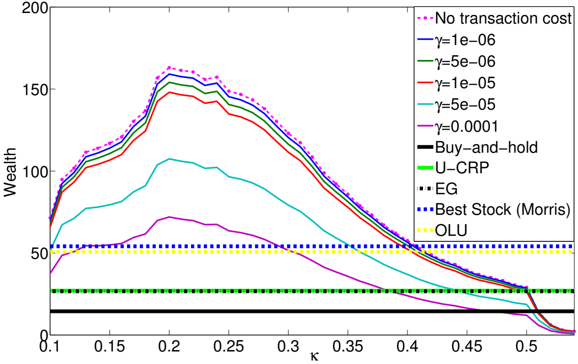
Takeaways:

- Higher η and $\kappa \Rightarrow$ more risk.
- Can effectively control risk via κ.

Transaction Cost-Adjusted Wealth

- γ: fixed percentage transaction cost.
- Competing algorithms:
  - Buy-and-hold: uniform investment, hold thereafter, i.e., no trading.
  - Uniform constant rebalanced portfolio (U-CRP): uniform investment, rebalance to uniform at end of each day.
  - Exponentiated Gradient (EG): multiplicative weight update strategy with KL-divergence as proximal term.
  - Online Lazy Updates (OLU): special case of ORASD with no groups or diversity constraint.

Transaction Cost-Adjusted Wealth



Takeaways:

- ORASD earns more than 3x wealth of best stock and competing algorithm even with transaction costs.
- Trade-off between risk and wealth is controlled by κ.

Conclusions

1. Introduced the $L_{(\infty,1)}$ group norm for structured diversification.
2. Presented an online resource allocation framework for diversification (ORASD) using $L_{(\infty,1)}$ as a constraint on the solution vector.
3. Presented an online portfolio selection algorithm and showed experimentally how risk is controlled via $L_{(\infty,1)}$ and κ and that ORASD earns more wealth than competing algorithms.

Conclusions

1. Introduced the $L_{(\infty,1)}$ group norm for structured diversification.
2. Presented an online resource allocation framework for diversification (ORASD) using $L_{(\infty,1)}$ as a constraint on the solution vector.
3. Presented an online portfolio selection algorithm and showed experimentally how risk is controlled via $L_{(\infty,1)}$ and κ and that ORASD earns more wealth than competing algorithms.

Thank you!

Conclusions

1. Introduced the $L_{(\infty,1)}$ group norm for structured diversification.
2. Presented an online resource allocation framework for diversification (ORASD) using $L_{(\infty,1)}$ as a constraint on the solution vector.
3. Presented an online portfolio selection algorithm and showed experimentally how risk is controlled via $L_{(\infty,1)}$ and κ and that ORASD earns more wealth than competing algorithms.

Thank you!

The research was supported in part by NSF grants IIS-1447566, IIS-1422557, CCF-1451986, CNS-1314560, IIS-0953274, IIS-1029711, and by NASA grant NNX12AQ39A. AB also acknowledges support from IBM and Yahoo. The authors would also like to thank Maria Gini for her helpful feedback.