On Tracking Portfolios with Certainty Equivalents on a generalization of Markowitz model: the Fool, the Wise and the Adaptive

RICHARD NOCK (1), BRICE MAGDALOU (1), ERIC BRIYS (1, 2) AND FRANK NIELSEN (3)
{rnock,bmagdalo}@martinique.univ-ag.fr, eric.b@cyberlibris.com, Nielsen@csl.sony.co.jp (1) CEREGMIA, Martinique, France, (2) Cyberlibris SA, Paris, France, (3) Sony CS Labs, Inc., Japan
Sony CSL
Sony Computer Science Laboratories, Inc

Portfolio allocation...

General investment framework

- *market*: d assets;
- *investor's portfolio* : allocation vector $\boldsymbol{\alpha} \in \mathbb{P}_d$ (d-dim. prob. simplex)
- *returns* : $\boldsymbol{w} \in [-1, +\infty)^d$, with $\boldsymbol{w} \sim p$ (density of returns)
- *investor's wealth* : average return $\omega_{\text{inv}} \doteq \boldsymbol{w}^\top \boldsymbol{\alpha}$

Investor: strategies to find $\boldsymbol{\alpha}$, to rank portfolios ($\boldsymbol{\alpha} \succeq \boldsymbol{\alpha}'$) ?

Strategies 0

BEHOLD, the fool saith, "Put not all thine eggs in the one basket"—which is but a manner of saying, "Scatter your money and your attention;" but the wise man saith, "Put all your eggs in the one basket and—WATCH THAT BASKET."—*Pudd'nhead Wilson's Calendar.*

Wise: $\boldsymbol{1}_i \succeq \boldsymbol{\alpha}$ (Mark Twain, *Pudd'nhead Wilson*, 1894)

Strategy I : (Non-trivial but) oversimplistic

$$\boldsymbol{\alpha} \succeq \boldsymbol{\alpha}' \Leftrightarrow \mathrm{E}_{\boldsymbol{w}\sim p}[\boldsymbol{w}^\top\boldsymbol{\alpha}] \geq \mathrm{E}_{\boldsymbol{w}\sim p}[\boldsymbol{w}^\top\boldsymbol{\alpha}']$$

Strategy II - Normative

I) Make five assumptions about the way the investor computes $\boldsymbol{\alpha} \succeq \boldsymbol{\alpha}'$: reflexivity, transitivity, completeness, continuity, separability.

Then,

$$\boldsymbol{\alpha} \succeq \boldsymbol{\alpha}' \Leftrightarrow \mathrm{E}_{\boldsymbol{w}\sim p}[\mathrm{U}(\boldsymbol{w}^\top\boldsymbol{\alpha})] \geq \mathrm{E}_{\boldsymbol{w}\sim p}[\mathrm{U}(\boldsymbol{w}^\top\boldsymbol{\alpha}')]$$

For some utility function U.

(Von Neumann & Morgenstern, 1944)

Risk premium and Certainty equivalent

Theorem:

parametrization of p

$$\mathrm{E}_{\boldsymbol{w}\sim p}[\mathrm{U}(\boldsymbol{w}^\top\boldsymbol{\alpha})] = \mathrm{U}(\underbrace{\mathrm{E}_{\boldsymbol{w}\sim p}[\boldsymbol{w}^\top\boldsymbol{\alpha}] - \mathrm{P}(\boldsymbol{\alpha};\boldsymbol{\theta})}_{\mathrm{C}(\boldsymbol{\alpha};\boldsymbol{\theta})})$$

$\mathrm{P}(\boldsymbol{\alpha};\boldsymbol{\theta})$ is the risk premium associated to $\boldsymbol{\alpha}$;

$\mathrm{C}(\boldsymbol{\alpha};\boldsymbol{\theta})$ is the certainty equivalent associated to $\boldsymbol{\alpha}$;

(Pratt, 1964)

$\mathrm{E}_{\boldsymbol{w}\sim p}[\mathrm{U}(\boldsymbol{w}^\top\boldsymbol{\alpha})]$ vs $\mathrm{U}(\mathrm{E}_{\boldsymbol{w}\sim p}[\boldsymbol{w}^\top\boldsymbol{\alpha}] - \mathrm{P}(\boldsymbol{\alpha};\boldsymbol{\theta}))$

Numerous cases — One case

... + Information Geometry

General information-geometric framework

- *exponential families*: $p(\boldsymbol{w}) = p_\psi(\boldsymbol{w}:\boldsymbol{\theta})$, with

$$\begin{aligned} p_\psi(\boldsymbol{w}:\boldsymbol{\theta}) &\doteq \exp\left(\boldsymbol{w}^\top\boldsymbol{\theta} - \psi(\boldsymbol{\theta})\right) b(\boldsymbol{w}) \\ &= \exp\left(-D_{\psi^\star}(\boldsymbol{w}\|\nabla_\psi(\boldsymbol{\theta})) + \psi^\star(\boldsymbol{w})\right) b(\boldsymbol{w}) \end{aligned}$$

$\boldsymbol{\theta}$ = natural parameters, $\psi : \mathbb{R}^d \rightarrow \mathbb{R}$ strictly convex differentiable
$\psi^\star(\boldsymbol{z}) \doteq \boldsymbol{z}^\top\nabla_\psi^{-1}(\boldsymbol{z}) - \psi(\nabla_\psi^{-1}(\boldsymbol{z}))$ convex conjugate of ψ

- *Bregman divergences*:

$$D_\psi(\boldsymbol{x}\|\boldsymbol{y}) \doteq \psi(\boldsymbol{x}) - \psi(\boldsymbol{y}) - (\boldsymbol{x}-\boldsymbol{y})^\top\nabla_\psi(\boldsymbol{y})$$

Deriving utility, risk premium and certainty equivalent in our information-geometric framework

Lemma: expression of U

Define

$$\mathrm{R}_i(\omega_{\text{inv}}) \doteq -\frac{\partial^2}{\partial w_i^2}\mathrm{U}(\omega_{\text{inv}})\left(\frac{\partial}{\partial w_i}\mathrm{U}(\omega_{\text{inv}})\right)^{-1}, \forall i = 1, 2, ..., d$$

as Arrow-Pratt measure of absolute risk aversion wrt stock i.

Assume Constant Absolute Risk Aversion (CARA) for some risk-aversion parameter $a \in \mathbb{R}$, *i.e.*: $\mathrm{R}_i(\omega_{\text{inv}}) = a, \forall i$.
Then

$$\mathrm{U}(x) = \begin{cases} x & \text{if } a = 0 \\ -\exp(-ax) & \text{otherwise} \end{cases}$$

Theorem: expression of $\mathrm{P}(\boldsymbol{\alpha};\boldsymbol{\theta}), \mathrm{C}(\boldsymbol{\alpha};\boldsymbol{\theta})$

Mean-divergence model

Assume CARA (a>0) and $p(\boldsymbol{w}) = p_\psi(\boldsymbol{w}:\boldsymbol{\theta})$.
Then

$$\begin{aligned} \mathrm{C}_\psi(\boldsymbol{\alpha};\boldsymbol{\theta}) &= \frac{1}{a}\left(\psi(\boldsymbol{\theta}) - \psi(\boldsymbol{\theta} - a\boldsymbol{\alpha})\right) \\ \mathrm{P}_\psi(\boldsymbol{\alpha};\boldsymbol{\theta}) &= \frac{1}{a}D_\psi\left(\boldsymbol{\theta} - a\boldsymbol{\alpha}\|\boldsymbol{\theta}\right) \end{aligned}$$

Particular case: Markowitz's risk premium (Markowitz, 1952) — Mean-variance model

$$p_\psi = \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \Rightarrow \mathrm{P}_{\boldsymbol{\mu},\boldsymbol{\Sigma}}(\boldsymbol{\alpha};\boldsymbol{\theta}) = \frac{a}{2}\boldsymbol{\alpha}^\top\boldsymbol{\Sigma}\boldsymbol{\alpha}$$

The investor cares "only" for mean and variance... but the Gaussian assumption does not hold in practice ! (Chavas, 2004)

From information geometry pops-up duality allocations vs returns

$\boldsymbol{w}$ $\quad \nabla_\psi^{-1} = \nabla_{\psi^\star} \quad \nabla_\psi \quad$ $\boldsymbol{\theta}$

$\boldsymbol{\theta}$ may be interpreted as a natural market allocation, information-theoretic optimal (generalizes a result by Markowitz, 1952, for Gaussians: $\boldsymbol{\alpha}_{\text{opt}} \propto \boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}$).

... On-line learning

General learning framework: on-line with three players

- the *natural market parameter* drifts without assumption:

$$\boldsymbol{\theta}_0, \boldsymbol{\theta}_1, \cdots$$

- a *reference portfolio* drifts without assumption:

$$\boldsymbol{r}_0, \boldsymbol{r}_1, \cdots$$

- we wish to update our portfolio,

$$\boldsymbol{\alpha}_0, \boldsymbol{\alpha}_1, \cdots$$

... so as to *lowerbound* our total certainty equivalent wrt reference:

$$\sum_{t=0}^{T-1}\mathrm{C}_\psi(\boldsymbol{\alpha}_t;\boldsymbol{\theta}_t) \geq \sum_{t=0}^{T-1}\mathrm{C}_\psi(\boldsymbol{r}_t;\boldsymbol{\theta}_t) - \text{loss}$$

(setting more general than assuming $\boldsymbol{r}_t = \boldsymbol{\theta}_t$, *e.g.* if $\boldsymbol{r}_t$ is good+sparse)

Algorithm: $\mathcal{OMD}_{\phi,\psi}$

- Initialize: initial portfolio $\boldsymbol{\alpha}_0 = (1/d)\boldsymbol{1}$, learning rate $\eta > 0$, strictly convex differentiable ϕ;
- Repeat for $t = 0, 1, ..., T-1$:

$$\boldsymbol{\alpha}_{t+1} \leftarrow \nabla_\phi^{-1}(\nabla_\phi(\boldsymbol{\alpha}_t) - \eta\nabla_{\mathrm{P}}(\boldsymbol{\alpha}_t;\boldsymbol{\theta}_t) - z_t\boldsymbol{1})$$

where z_t is chosen s.t. $\boldsymbol{\alpha}_{t+1} \in \mathbb{P}$ (+scale & renormalize if $\mathrm{dom}(\phi) \not\subset \mathbb{R}_+$)

Properties of $\mathcal{OMD}_{\phi,\psi}$

Theorem: lowerbound on certainty equivalents for $\mathcal{OMD}_{\mathrm{KL},\psi}$

Fix $a = K/\min_{t\in\mathbb{T}}\|\boldsymbol{\alpha}_t - \boldsymbol{r}_t\|_2^2$. Then, for any $\eta > 0$,

$$\sum_{t=0}^{T-1}\mathrm{C}_\psi(\boldsymbol{\alpha}_t;\boldsymbol{\theta}_t) \geq \sum_{t=0}^{T-1}\mathrm{C}_\psi(\boldsymbol{r}_t;\boldsymbol{\theta}_t) - d^{\frac{1}{q}}K'\sum_{t=0}^{T-1}\|\boldsymbol{r}_{t+1} - \boldsymbol{r}_t\|_p - K''$$

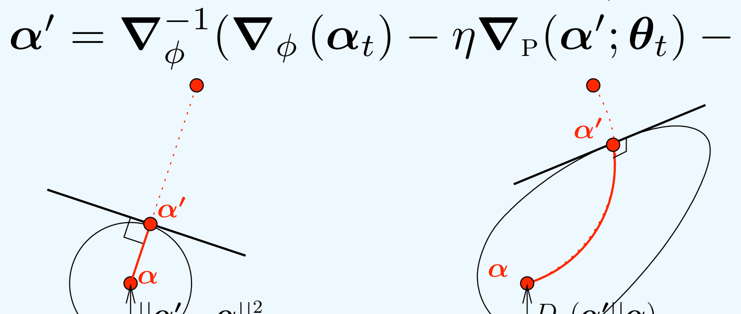
(see the paper for the expressions of K, K', K''; $p, q \geq 1, (1/p) + (1/q) = 1$)

Lemma: $\mathcal{OMD}_{\phi,\psi}$ uses a generalization of Amari's natural gradient

The solution to $\boldsymbol{\alpha}' = \arg\min_{\boldsymbol{\alpha}\in\mathbb{A}} D_\phi(\boldsymbol{\alpha}\|\boldsymbol{\alpha}_t)$

$$\mathbb{A} = \{\boldsymbol{\alpha} \in \mathbb{R} : (\boldsymbol{\alpha}^\top\boldsymbol{1} = 1) \wedge (\mathrm{P}_\psi(\boldsymbol{\alpha};\boldsymbol{\theta}) \leq k)\}$$

satisfies: $\boldsymbol{\alpha}' = \nabla_\phi^{-1}(\nabla_\phi(\boldsymbol{\alpha}_t) - \eta\nabla_{\mathrm{P}}(\boldsymbol{\alpha}';\boldsymbol{\theta}_t) - z_t\boldsymbol{1})$



(Amari, 1998)

Experiments

General experimental framework

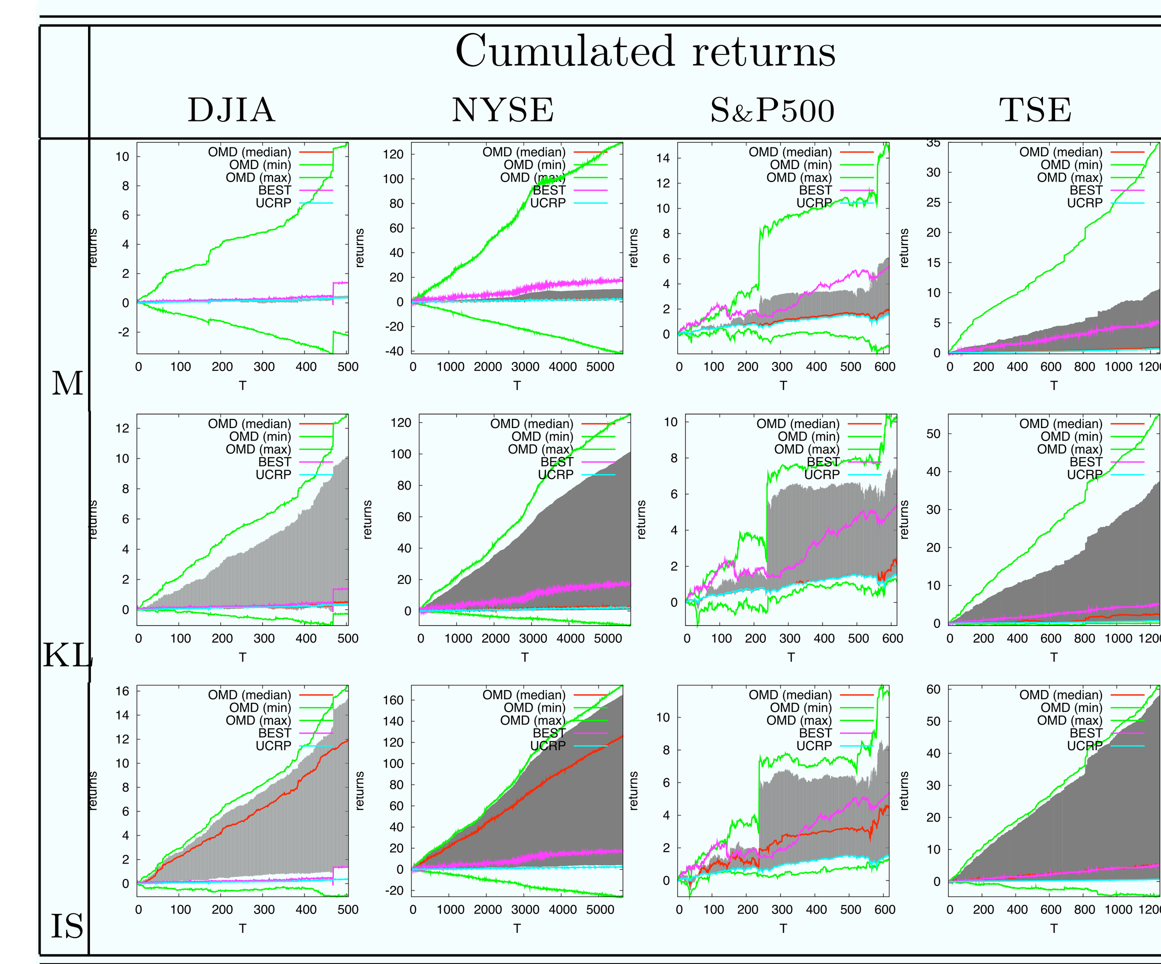
- 4 markets with daily (DJIA, NYSE, TSE) or weekly (S&P500) returns:

name	d	T	start date	end date
DJIA	30	506	01/14/01	01/14/03
NYSE	36	5650	07/03/62	12/31/84
S&P500	324	618	01/08/98	11/12/09
TSE	88	1257	01/04/94	12/31/98

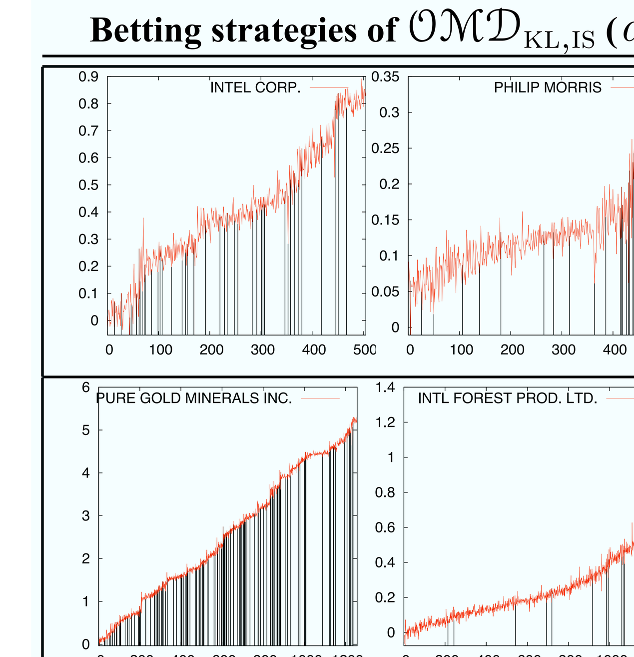
- Contenders: UCRP (Uniform Cost Rebalanced Portfolio) & Best stock
- Parameters (more tests in the supplementary material)

$a \in \{0.01, 1, 100\}, \eta \in \{0.01, 1, 100\}, \psi \in \{\mathrm{M}, \mathrm{KL}, \mathrm{IS}\}$

M $\Rightarrow D_\psi$ = Mahalanobis ($\mathrm{P}_\psi(\boldsymbol{\alpha};\boldsymbol{\theta})$ = Markowitz' premium)
KL $\Rightarrow D_\psi$ = Kullback-Leibler
IS $\Rightarrow D_\psi$ = Itakura-Saito



Betting strategies of $\mathcal{OMD}_{\mathrm{KL,IS}}$ ($a = 100.0, \eta = 0.01$) :



DJIA::

None:	16.01%
INTEL CORP.:	7.91%
(⋆) AT&T CORP.:	7.31%
HP:	6.32%
JP MORGAN:	4.94%
PHILIP MORRIS:	4.35%
HONEYWELL:	4.35%

TSE:

None::	17.33%
(⋆) PURE GOLD MIN.:	9.70%
BREAKWATER RES.:	8.27%
REPAP ENT. INC.:	5.72%
GENTRA INC.:	3.50%
COTT CORP.:	3.34%
MIRAMAR MIN.:	3.18%

Left: vertical bars indicate when stock had absolute majority in the portfolio;
Right: % of iterations during which one stock (or None) had absolute majority.

Work supported by ANR / Additional material (proofs, experiments) available at: *http://www1.univ-ag.fr/~rnock/Articles/ICML11/*