Exploring non linearities in Hedge Funds

An application of Particle Filters to Hedge Fund Replication

Guillaume Weisang[1] Thierry Roncalli[2]

[1]Bentley University, Waltham, MA

[2]Lyxor AM

January 28-29, 2010

Overview of the Factor Approach

A hedge fund portfolio: $\boxed{r_t^{\text{HF}} = \sum_{i \in I} \omega_{it} r_{it}}$

Assumption

The structure of all asset returns can be summarized by a set of risk factors $\{F_j\}_{j=1,\ldots,m}$:

$$\forall t \qquad r_{it} = \alpha_i + \sum_{j=1}^{m} \beta_{ij} F_{jt} + \xi_{it}$$

with

$$\mathbb{E}\left(\xi_{it} \left| F_{1t} \ldots F_{mt} \right.\right) = 0$$

A typical factor model

One assumes

$$\boxed{r_t^{\text{HF}} = \alpha_t^{\text{HF}} + \sum_{j=1}^{m} w_{jt}^{\text{HF}} F_{jt} + \varepsilon_t}$$

such that

$$\begin{aligned} \alpha_t^{\text{HF}} &= \sum_{i \in I} \alpha_i \omega_{it} \\ \boxed{\text{risk exposures}} \longrightarrow w_{jt}^{\text{HF}} &= \sum_{i \in I} \omega_{it} \beta_{ij} \\ \varepsilon_t &= \sum_{i \in I} \omega_{it} \xi_{it} \end{aligned}$$

Literature Review of the Factor Approach

Summary

- Static Linear factor models [Amenc et al., 2007]
  - Lack reactivity
  - Fail the test of robustness, giving poor out-of-sample results
- Factor selection [Fung and Hsieh, 1997] [Lo, 2008]
  - In static models, **economic selection** of factors ⟶ significant improvement over other methodologies for out-of-sample robustness test.
  - In dynamic models, [Darolles and Mero, 2007] uses a PCA-based factor evaluation methodology [Bai and Ng, 2006] on rolling OLS regressions.
    - ★ Improvement over “naive” inclusion of all relevant economic factors
    - ★ Poor Interpretability of the evaluated factors
- Dynamic linear models [Roncalli and Teiletche, 2008] [Lo, 2008] [Jaeger, 2009]:
  Capturing the *unobservable* dynamic allocation using **traditional (OLS) methods** is
  - Very difficult
  - Estimates can vary greatly at balancing dates
- Nonlinear models $\rightleftharpoons$ methodological challenge [Amenc et al., 2008] [Diez de los Rios and Garcia, 2008]

Hedge Fund Replication – The Nonlinear Non-Gaussian Case

Why It is Interesting

- HF Returns are not Gaussian
  - negative skewness and positive excess kurtosis.
- Nonlinearities in HF Returns
  - Nonlinearities documented from the very start of hedge-fund replication – see, e.g., [Fung and Hsieh, 1997].
  - Nonlinearities
    - are important for some strategies but not for the entire industry [Diez de los Rios and Garcia, 2008].
    - may be due to positions in derivative instruments or un-captured dynamic strategies – see, e.g., [Merton, 1981].
  - No successful hedge fund replication using non-linear models has ever been done

Tracking Problems

Definition (Tracking Problem)

The following two equations define a tracking problem (TP) [Arulampalam et al., 2002]:

$$\left\{\begin{array}{lcll} \mathbf{x}_k & = & f(t_k, \mathbf{x}_{k-1}, v_k) & \text{(Transition Equation)} \\ \underset{\uparrow}{\mathbf{z}_k} & = & h(t_k, \underset{\uparrow}{\mathbf{x}_k}, \eta_k) & \text{(Measurement Equation)} \end{array}\right.$$

where “shadow” object

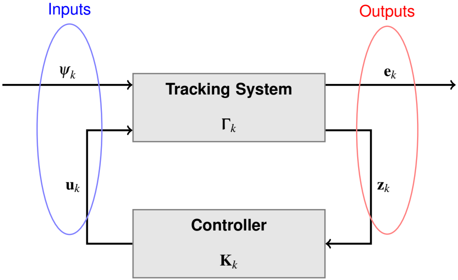
- $\mathbf{x}_k \in \mathbb{R}^{n_x}$ is the state vector, and $\mathbf{z}_k \in \mathbb{R}^{n_z}$ the measurement vector at step k.
- v_k et η_k are mutually independent i.i.d noise processes.
- The functions f and h can be non-linear functions.

Tracking Problems and Tactical Allocation

Tracking Systems

Discrete case, at time step k

HFR: $\mathbf{x}_k = (w_{1k}^{\text{HF}}, \ldots, w_{mk}^{\text{HF}})'$ ⇐ risk exposures



- **Outputs**
  - Tracking Error
    $\mathbf{e}_k = \mathbf{z}_k - \hat{\mathbf{z}}_{k|k-1}$
    $\mathbf{e}_k = r_k^{\text{HF}} - r_k^{\text{Clone}}$
  - Censored measurement $\mathbf{z}_k$
    $\mathbf{z}_k = r_k^{\text{HF}}$
- **Inputs**
  - Exogenous signals
    $\psi_k = (\mathbf{x_0}, \eta_{1:k}, \nu_{1:k})$
    HF changes in allocation, strategies or reporting
  - Controlled input $\mathbf{u}_k$
    Assumption:
    $\mathbf{u}_k = K_k \mathbf{z}_k$
    Adjustments to the replication portfolio's risk exposures

Bayesian Filters

Optimal Control Theory

Under some general assumptions, one can prove

$$\text{tracking error} \rightarrow \mathbf{e}_k = \mathbf{T}_{e\psi_k} \; \psi_k \leftarrow \text{exogenous signals}$$

with

$$\text{transfer function} \rightarrow \mathbf{T}_{e\psi_k} = \Gamma_{e\psi_k} + \Gamma_{eu_k} \; K_k \; \left(I - \Gamma_{zu_k}\right)^{-1} \Gamma_{z\psi_k}$$

The role of the controller K_k is to

- stabilize the system
- make $\mathbf{T}_{e\psi}$ small in an appropriate sense.

Definition (Stability)

A system is said to be marginally stable if the state $\mathbf{x}$ is bounded for all time t and for all bounded initial states $\mathbf{x}_0$.

Bayesian Filters are algorithms which provide the **optimal estimators** of the state $\mathbf{x}_k$

Advantages of Bayesian Filters: no assumption of stationarity

Bayesian Filters

Solving Tracking Problems

At time step k

- Prediction equation for the prior density

$$p\left(\mathbf{x}_k \mid \mathbf{z}_{1:k-1}\right) = \int p\left(\mathbf{x}_k \mid \mathbf{x}_{k-1}\right) p\left(\mathbf{x}_{k-1} \mid \mathbf{z}_{1:k-1}\right) \mathrm{d}\mathbf{x}_{k-1}$$

- Update equation for the posterior density

$$p\left(\mathbf{x}_k \mid \mathbf{z}_{1:k}\right) \propto p\left(\mathbf{z}_k \mid \mathbf{x}_k\right) p\left(\mathbf{x}_k \mid \mathbf{z}_{1:k-1}\right)$$

- Best estimates

$$\hat{\mathbf{x}}_{k|k-1} = \mathbb{E}\left[\mathbf{x}_k \mid \mathbf{z}_{1:k-1}\right] \qquad \hat{\mathbf{x}}_{k|k} = \mathbb{E}\left[\mathbf{x}_k \mid \mathbf{z}_{1:k}\right]$$

- Implementation
  - Kalman Filter (KF): linear Gaussian case
  - H_∞ Filters or Particle Filters (PF): nonlinear or non Gaussian case

Particle Filters

- If the posterior density $p\left(\mathbf{x}_k \mid \mathbf{z}_{1:k}\right) \propto \pi\left(\mathbf{x}_k\right)$ such that $\pi\left(x\right)$ is easy to evaluate but difficult to draw sample from.
- Let $\left\{\mathbf{x}_k^i\right\}_{i=1}^{N_s}$ be samples from an importance density $q\left(\cdot\right)$
- The posterior density at time k can then be approximated as

$$p\left(\mathbf{x}_k \mid \mathbf{z}_{1:k}\right) \approx \sum_{i=1}^{N_s} w_k^i \delta\left(\mathbf{x}_k - \mathbf{x}_k^i\right) \tag{1}$$

where, using **Bayes rule**,

$$\boxed{w_k^i \propto w_{k-1}^i \frac{p\left(\mathbf{z}_k \mid \mathbf{x}_k^i\right) \times p\left(\mathbf{x}_k^i \mid \mathbf{x}_{1:k-1}^i\right)}{q\left(\mathbf{x}_k^i \mid \mathbf{x}_{1:k-1}^i, \mathbf{z}_k\right)}}$$

- The set of support points $\left\{\mathbf{x}_k^i\right\}_{i=1}^{N_s}$ and their associated weights $\left\{w_k^i, i = 1, \ldots, N_s\right\}$ characterizes the *posterior density* at time step k
- Equation (1) is at the core of Particle Filters (PF). Considering different assumptions leads to different numerical algorithms (SIS, GPF, SIR, RPF, etc.).

Hedge Fund Replication – Non-Gaussian Nonlinear Case

Objectives

Non-Gaussian or Nonlinear Case

Non Gaussian

$$\left\{\begin{array}{lcl} \mathbf{w}_k^{\text{HF}} & = & \mathbf{w}_{k-1}^{\text{HF}} + \nu_k \\ r_k^{\text{HF}} & = & \mathbf{r}_k' \mathbf{w}_k^{\text{HF}} + \eta_k \\ \eta_k & \sim & \mathscr{H} \end{array}\right.$$

with $\mathscr{H}$ non Gaussian

Nonlinear

$$\left\{\begin{array}{lcl} \mathbf{w}_k^{\text{HF}} & = & \mathbf{w}_{k-1}^{\text{HF}} + \nu_k \\ r_k^{\text{HF}} & = & \mathbf{r}_k' \mathbf{w}_k^{\text{HF}} \\ & + & w_{k,(m+1)}^{\text{HF}} r_{k,(m+1)}\left(s_k\right) + \eta_k \end{array}\right.$$

with $r_{k,(m+1)}\left(s_k\right)$ nonlinear

e.g., option on S&P 500

The objectives of this paper are to

- explore the nature of HF nonlinearities
    1. non Gaussian errors
    2. nonlinear factor
- explore possible remedy for Hedge Fund replication: PF

The Gaussian Distribution Assumption

Framework

- Consider

$$\begin{cases} r_k^{\mathrm{HF}} = \mathbf{r}_k^\top \beta_k + \eta_k \\ \beta_k = \beta_{k-1} + \nu_k \\ \eta_k \sim \mathscr{H} \end{cases}$$

with $\mathscr{H}$ non Gaussian $\longrightarrow$ May be solved using Particle Filters.

- Assume $\mathscr{H}$ is a Skew t distribution $\mathscr{S}\left(\mu_\eta, \sigma_\eta, \alpha_\eta, \nu_\eta\right)$
- 3 estimation methods

(PF #1) ML on parameters of $\mathscr{H}$ using the Kalman Filter (KF) tracking errors.

(PF #2) GMM to estimate $m+3$ parameters (classical MM + two moments for skewness and kurtosis).

(PF #3) Same as (PF #2) except $\hat{\alpha}_\eta$ is forced to -10.

The Gaussian Distribution Assumption

Results with SIR algorithm and 50000 particles

	$\hat{\mu}_{\mathrm{1Y}}$	$\hat{\sigma}_{\mathrm{1Y}}$	s	γ_1	γ_2
HF	9.94	7.06	0.77	-0.57	2.76
LKF[1]	7.55	6.91	0.45	-0.02	2.25
PF #1	7.76	7.44	0.45	-0.03	2.02
PF #2	7.57	7.28	0.43	-0.11	1.93
PF #3	6.90	7.99	0.31	-0.57	2.88

	π_{AB}	σ_{TE}	ρ	τ	ρ_S
LKF	75.93	3.52	87.35	67.10	84.96
PF #1	78.09	4.03	84.71	63.49	81.94
PF #2	76.13	4.25	82.51	61.60	80.20
PF #3	69.43	5.11	77.62	54.75	73.55

Conclusion

With **linear assets**, higher kurtosis and negative skewness come at the cost of a higher tracking error σ_{TE}.

⇒ It is not the right way to get at nonlinearities.

---

[1] LKF = linear 6F model estimated using Kalman filter.

Taking into account Nonlinear Assets

Idea

Build Option Factors

$$\begin{cases} \mathbf{w}_k^{\text{HF}} & = & \mathbf{w}_{k-1}^{\text{HF}} + \nu_k \\ r_k^{\text{HF}} & = & \mathbf{r}_k' \mathbf{w}_k^{\text{HF}} \\ & + & w_{k-1,(m+1)}^{\text{HF}} r_{k,(m+1)}(s_k) + \eta_k \end{cases}$$

where

- $r_{k,(m+1)}(s_k)$ nonlinear, (e.g., the return of a systematic one-month option selling strategy on S&P 500)
- s_k is the strike of the option at time index k.

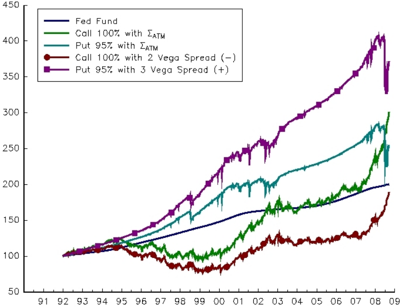
Exogenous Option Factors

Problem

Results are data dependent: liquidity, bid/ask spread, size amount (e.g., backtest with VIX).

Example

Systematic selling at end of month 1M put (resp. call) options with $s_k = 95\%$ (respectively 100%).



Conclusion

Dependent on

- Rebalancing dates (e.g., end of month certainly a most favorable time for selling put options).
- Implied volatility data and on skew's and bid/ask spread's assumptions.

Endogenous Option Factors

Estimation procedure

Two ways to estimate these strikes

(a) Estimate the option strikes separately from the tracking problem: endogenous to the estimation but exogenous to the filter (which can then use KF).

(b) The option strike belongs to the state vector of a nonlinear TP system

$$\left\{ \begin{array}{l} \begin{pmatrix} \mathbf{w}_k \\ s_k \end{pmatrix} = \begin{pmatrix} \mathbf{w}_{k-1} \\ s_{k-1} \end{pmatrix} + \begin{pmatrix} \nu_k \\ \varepsilon_k \end{pmatrix} \\ r_k^{(\mathrm{HF})} = \sum_{i=1}^{m} w_k^{(i)} r_k^{(i)} + w_k^{(m+1)} r_k^{(m+1)} (s_k) + \eta_k \end{array} \right. \tag{2}$$

- PF
- $\eta_k \sim \mathcal{N}\left(0, \sigma_\eta^2\right)$ and

$$\left(\begin{array}{c} \nu_k \\ \varepsilon_k \end{array} \right) \sim \mathcal{N}\left(\left(\begin{array}{c} \mathbf{0} \\ 0 \end{array} \right), \left(\begin{array}{cc} Q & 0 \\ 0 & \sigma_s^2 \end{array} \right) \right)$$

with $Q = \operatorname{diag}\left(\sigma_1^2, \ldots, \sigma_m, \sigma_{m+1}^2\right)$.
The vector of unkown parameters to estimate is then $\theta = \left\{\sigma_1, \ldots, \sigma_m, \sigma_{m+1}, \sigma_s, \sigma_\eta\right\}$.

Endogenous Option Factors

Estimation procedure (Cont'd)

- Direct estimation of system (2)'s parameters by PF and (relative) scarcity of data on HF returns is very difficult
  $\Longrightarrow$ use of yet another estimation method

Step 1 σ_η and σ_i for $i = 1, \ldots, m$: ML considering the linear factor model

Step 2 σ_{m+1} and σ_s: grid-based method *conditionally on the previous estimates.*

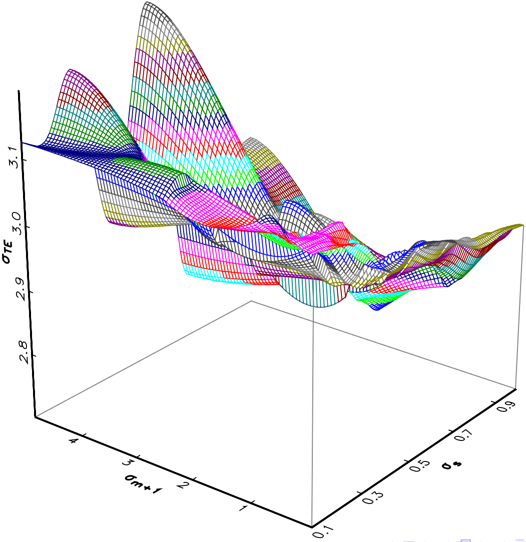
If f denotes the statistic of interest in the maximization (or minimization) of and if Ω denotes the set of grid points:

$$\{\hat{\sigma}_{m+1}, \hat{\sigma}_s\} = \arg\max f\left(\sigma_{m+1}, \sigma_s \mid \hat{\sigma}_1, \ldots, \hat{\sigma}_m, \hat{\sigma}_\eta\right) \quad \text{u.c.} \ \left(\sigma_{m+1}, \sigma_s\right) \in \Omega.$$

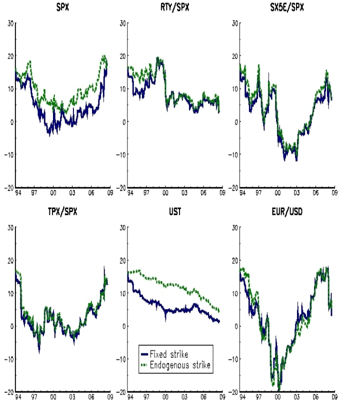
Results are biased yet consistent w.r.t. linear model

Endogenous Option Factors

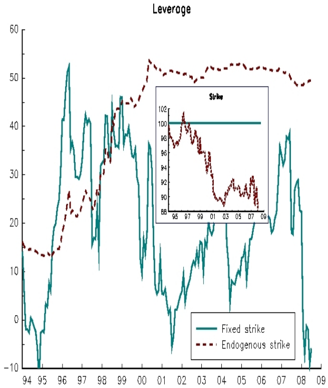
Example: Grid approach applied to the HFRI RV index



Endogenous Option Factors



(a) Exposures of the linear assets for the HFRI RV index



(b) Option exposures and strikes for the HFRI RV index

HFR – Non-Gaussian Nonlinear Case

Key points and Future Developments

Gaussian assumption KF's tracking errors have skew and excess kurtosis. A remedy: Skew t distribution

1. very difficult direct estimation of parameters in PF
2. no luck with two-step procedure (KF + GMM) $\Longrightarrow \searrow$ Skew, $\nearrow$ TE

Nonlinear Factor Endogenous and exogenous

1. Exogenous factors are extremely data dependent
2. Endogenous factors: some success using a grid-based approach and KF; PF code has to be parallelized
3. For now, purely academic exercise

Robust Methodology

1. TO BE DEVELOPED
2. H_∞ Filters minimize worst cases $\longrightarrow$ robust to violations of Gaussian and linearity assumptions

Selected References I

Thierry Roncalli and Guillaume Weisang
Tracking Problems, Hedge Fund Replication and Alternative Beta.
Working Paper, available on SSRN, 2009.

Andrew Lo
Hedge Funds: An Analytic Perspective.
Princeton University Press, 2008.

N. Amenc, W. Géhin, L. Martellini, and J-C. Meyfredi.
The Myths and Limits of Passive Hedge Fund Replication.
Working Paper, June 2007.

N. Amenc, L. Martellini, J-C. Meyfredi and V. Ziemann.
Passive Hedge Fund Replication – Beyong the Linear Case.
Working Paper, January 2008.

S. Darolles and G. Mero.
Hedge Fund Replication and Factor Models.
Working Paper, 2007.

Selected References II

J. Bai and S. Ng.
Evaluating Latent and Observed Factors in Macroeconomics and Finance.
Journal of Econometrics, 131(1-2):507-537, 2006.

A. Diez de los Rios and R. Garcia.
Assessing and Valuing the Non-Linear Structure of Hedge Fund Returns.
Working Paper, 2008.

W. Fung and D. A. Hsieh.
Empirical Characteristics of Dynamic Trading Strategies: the Case of Hedge Funds.
Review of Financial Studies, 10:275-302, 1997.

T. Roncalli and J. Teiletche.
An Alternative Approach to Alternative Beta.
Journal of Financial Transformation, 24:43-52, 2008.
Available at SSRN: `http://ssrn.com/abstract=1035521`.

Selected References III

R. C. Merton.
On Market Timing and Investment Performance. I. An Equilibrium Theory of Value for Market Forecasts.
Journal of Business, 54(3):363-406, 1981.

S. Arulampalam, S. Maskell, N.J. Gordon and T. Clapp.
A Tutorial on Particle Filters for Online nonlinear/non-Gaussian Bayesian Tracking.

IEEE Transaction on Signal Processing, 50(2):174-188, Februrary 2002.

Lars Jaeger.
Alternative Beta Strategies and Hedge Fund Replication.
John Wiley & Sons, NY, 1^{st} edition, 2008.

Statistics Description

- $\hat{\mu}_{1\mathrm{Y}}$ is the annualized performance;
- π_{AB} the proportion of the HFRI index performance explained by the clone;
- σ_{TE} is the yearly tracking error;
- ρ, τ and ρ_S are respectively the linear correlation, the Kendall tau and the Spearman rho between the monthly returns of the clone and the HFRI index;
- s is the sharpe ratio;
- γ_1 is the skewness;
- γ_2 is the excess kurtosis.