Dominant Factor™ Analysis by Nonlinear Polymodels

RAPHAEL DOUADY
Stony Brook University
Labex ReFi, CNRS

OCTOBER 9, 2017
MCFAM DISTINGUISHED LECTURE
Traditional Multi-Factor Analysis

\( Y = \) Dependent random variable

\( X_1, \ldots, X_n = \) Independent variables

Explanation: \( Y = f(X_1, \ldots, X_n) + \varepsilon \)

Ideally: \( f(X_1, \ldots, X_n) = E[Y|X_1, \ldots, X_n] \)

Statistician’s questions:

- Distribution of residuals \( \varepsilon \)
- Quality of the fit, spurious results?
- Convergence of estimators of \( f \)
- Relevance of variables, missing independent variables?

Most often, when \( n > 1 \), one chooses linear \( f \) \( \Rightarrow \) Poor prediction power
Practitioner’s Questions

- What are the sources of risk? Which scenarios are dangerous?
  - Which variables $X_i$ have an influence on $Y$ when it moves dangerously?
  - How far can they go?

- How far can it go? How frequently? How much should I pay (or charge) for an insurance?
  - What is the shape of the return distribution?
  - Fat tails?

- Can I anticipate dangerous events? Can I protect myself against?
  - Crisis prediction (Type I and Type II errors)
  - Hedge efficiency (nature of the hedge, insurer’s quality, liquidity risk...)

- Diversification
  - Portfolio return distribution
  - Contribution of each component to the overall risk
  - Portfolio optimization
Problems with Multi-Factor Analysis

- Regime changes invalidate calibration
  - Model $f_t$ calibrated on $[t - \tau, t]$
  - Conditional prediction $\hat{Y}_{t+1} = f_t(X_{1,t+1}, ..., X_{n,t+1})$
  - Out-of-sample prediction power: $P-Square = 1 - \frac{\text{Var}(Y_{t+1} - \hat{Y}_{t+1})}{\text{Var}(Y_{t+1})}$ often negative!
  - Reasons: model rigidity (linearity), spurious calibration (overfitting), fitting instability (collinearity)

- “Useful most of the time... except when needed!”
  - Question is when is the model being used $\Rightarrow$ Calibration conditionally to usage (e.g. LOWESS)
  - Problem: usage conditions are rare events

- Vanishing Diversification
  - Correlations depend on the regime
  - Optimized portfolios have fatter tails than non-optimized ones
Optimizers Failed, However Advanced...

- Optimizer
  - MAX: Expected Return
  - MIN: Risk

- Real Risk
  - MEASURED RISK
  - HIDDEN RISK

- Optimization
  - Maximizes the Ratio of Hidden/Measured Risk
Vanishing Diversification

Optimizers, however sophisticated, simply maximize expected return while minimizing measured risk. Therefore, by design, optimizers maximize the proportion of unmeasurable risk – i.e. hidden risk – leading automatically to portfolios which eventually deliver very nasty surprises.

Assume that $Y_1...Y_m$ have mixed joint distribution $P = \pi_1 P_1 +...+ \pi_q P_q$ with $\pi_1 >...> \pi_q$

Fat tails can be measured as the ratio of risk under $P_1$ vs. risk under other regimes.

An optimizer that only accounts for some regimes will reduce the risk under those regimes, but increase the risk under other regimes, hence increasing fat-tailedness

$\Rightarrow$ Regime changes have an aggravated impact on portfolio risk
The Data Wall

- Millions of Assets and Funds
  - A few years of history => only a few 10’s to a few 100’s of returns
  - Regime changes: future ≠ past
  - Position info: unreliable, incomplete, delayed, fast changing
  - Large variety of strategies and trading universe

- 10,000’s Risk Factors
  - All asset classes
  - Long term history, including many crises, cycles
  - Hedge Funds often uncorrelated to markets: need exotic factors
  - Correlations only appear during crises: need nonlinear models

Too many models, too little information

IMPOSSIBLE TO SELECT AND CALIBRATE A MODEL
What Are You Looking For?

Did you lose your key there?

No, on the other side, but here I have light!
Polymodel Principle

- Replace one function \( Y = f(X_1, \ldots, X_n) + \varepsilon \) by a collection:

\[
\begin{align*}
Y &= f_1(X_1) + \varepsilon_1 \\
&= \mathbb{E}[Y | X_1] \\
&
\vdots \\
Y &= f_n(X_n) + \varepsilon_n \\
&= \mathbb{E}[Y | X_n]
\end{align*}
\]

Taken individually, each model is imprecise, but the collection contains at least as much information as a multi-factor model, and in fact much more!

- Individual models can be nonlinear and contain lags
  - Easy to calibrate
  - Focus on where useful info lies: in the extremes

- Factors can be correlated and linearly, or even nonlinearly dependent
  - Use hundreds, or even thousands of factors
  - Need to assess the relevance of each factor \( \Rightarrow p\text{-value} \)
This performance series would attract any investor who is solely focused on past performances. The sequel shows how the investor might be disappointed.
Pure Performance Analysis

Could such a loss be anticipated, only looking at past fund performance?

Yes, with nonlinear factor analysis

Credit driven fund:

• Long AAA bonds, Short T-bonds, duration 10Y

Sharpe = -0.25
Annualised Volatility = 3.4%
Annualised return = 2.6%
VaR 99 = 3.5% (3.5 sigma)
Peak to valley = 12.2%
Skew = -1.0
Excess Kurtosis = 3.0
Single Factor Analysis

These fund returns depend mostly on the AAA credit spread – in a nonlinear (optional) manner.

The grey curve is obtained by aggregating the nonlinear function of credit spread changes over many years.

This leads us to a novel approach for anticipating extreme risk, namely through the concept of STRESS VAR.

Credit driven fund vs. AAA spread over T-Bonds:

- This fund was just surfing the good wave during the analysis period
One can see that the loss experienced in 2007 had several similar precedents. The loss of the fund is in line with its Stress VaR, which itself is derived from “extrapolated” losses of the fund, prior to its actual track record.

> Credit driven fund vs. AAA spread over T-Bonds:
  - The driving factor experienced many jumps in the past comparable to the crisis
Risk Measurement by Polymodels

**Step 1:** Identify a **LARGE** Set of Factors
- **LONG HISTORY** (20 Yrs incl. crises)
- As many factors as potential risk sources ⇒ Several 100’s

**Step 2:** Scan Factors One at a Time
- Select “dominant factors”™ with a strong statistical relationship to the fund ⇒ p-value $p_i$
- E.g. Reject factors with $p_i > \min(p_j, j = 1...n)^{\alpha}$ \quad $\alpha < 1$
- Focus on **EXTREME MOVES** ⇒ Nonlinear Models

**Step 3:** Stress Dominant Factors™
- Worst Impact of each dominant factor™
- Single Factor Stress VaR = $\sqrt{(WORST IMPACT)^2 + SPECIFIC^2}$
- STRESS VaR = MAX(Single Factor Stress VaR)

Poly-models are aimed at breaking the “data wall”. Here, the major innovation is in the way that the distribution of future returns is estimated; using a very long history of markets in order to include past crises, a large number of factors in order to account for all possible risk sources and a collection of nonlinear models in order to account for extreme risks – in particular, the impact of liquidity gaps. Limited fund historical records are utilized in an optimal way.
Risk Measurement by Polymodels

Handle 100’s of Market Factors

Model rare events (“Black Swans”)

More accurate when needed, than when not needed!

- Tail concentration effect

Suited for risk measurement and stress scenarios

- Prediction from individual factors can be merged
- Risk measure = STRESS VaR (worst case) includes hidden risks

Can be aggregated for a portfolio

- Risk contributions involve extreme correlations
- Superior allocation and optimization
Relation between Multi-Factor and Polymodels

Linear Multi-Factor Model

\[ Y = \lambda_1 X_1 + \ldots + \lambda_n X_n + \varepsilon \]

- Coefficient \( \lambda_i \) are assumed fixed
- The Factor set \( \{X_1, \ldots, X_n\} \) is frozen

Poly-Model:

- Linear:
  \[ Y = \beta_i X_i + \varepsilon_i \quad i = 1 \ldots n \]

- Nonlinear + lags:
  \[ Y = \phi_i(X_i) + \psi_i(X_i(t-1)) + \rho_i X(t-1) + \varepsilon_i \quad i = 1 \ldots n \]

- Score each model by relevance in extreme scenarios
- \( p_i = p\text{-value} \) of model with \( X_i \) computed by overweighting large returns
Relation between Multi-Factor and Polymodels

Relation with Multi-factor Models: the **Linear case**

- \( Y = \beta_i X_i + \alpha_i \quad i = 1\ldots n \)
- \( Y = \lambda_1 X_1 + \ldots + \lambda_n X_n + \alpha \)
- \( <Y, X_i> = \beta_i \text{Var}(X_i) = \sum \lambda_j <X_i, X_j> \)
- \( (\lambda_1,\ldots, \lambda_n) = \text{Cov}(X)^{-1} (\beta_1 \sigma_1^2, \ldots, \beta_n \sigma_n^2) \)
- The uncertainty on \( \lambda_i \)'s depends on colinearity of factors
- Badly conditioned covariance matrix \( \Rightarrow \) Low Information Ratio

**Nonlinear Modelling**

- Decompose each \( \varphi_i \) over Hermitte Polynomials \( H_k \):
  \[ \varphi_i(X_i) = \sum \beta_i^k H_k(X_i) + \alpha_i \]
- Nonlinear Multi-factor model by inverting \( \text{Cov}(H_k(X_i)) \)
- We obtain: \( Y = \psi_1(X_1) + \ldots + \psi_n(X_n) + \varepsilon \)
- Improve Information Ratio with LOWESS Regression
Relation between Multi-Factor and Polymodels

Theorem (Cherny-Douady)

Given $X_1, \ldots, X_n$ and functions $\varphi_1, \ldots, \varphi_n$ such that $E[\varphi_1(X_1)] = \ldots = E[\varphi_n(X_n)]$

Does there exist a function $\varphi$ of $n$ variables such that:

$$\forall i \quad E[\varphi(X_1, \ldots, X_n) \mid X_i] = \varphi_i(X_i)$$

Answer: the minimum variance solution is given by

$$\varphi(X_1, \ldots, X_n) = \psi_1(X_1) + \ldots + \psi_n(X_n)$$

It exists provided an ellipticity condition is satisfied: for any $(\varphi_1, \ldots, \varphi_n) \in L^2(X_1) \times \ldots \times L^2(X_n)$

$$\left\| \sum_{i=1}^n \varphi_i(X_i) \right\|^2_{L^2} \geq c \sum_{i=1}^n \left\| \varphi_i(X_i) \right\|^2_{L^2}$$
Stress Testing by Poly-Models

Information Ratio

- Given $I = (i_1, ..., i_q)$ and factor stress values $(x_{i1}, ..., x_{iq})$ we compute the joint impact by merging single factor models:

$$\text{Impact} = \sum_{i \in I} \lambda_{ik} H_k(x_i) + \alpha_I$$

where $\lambda_{ik}$ are the coefficients of the merged multi-factor nonlinear model.

- The uncertainty of the estimate is given by the covariance matrix of coefficients $(\lambda_{ik}, \alpha_I)$, which can be redeemed from the inverse Hessian of the log-likelihood function.

$$\text{Information Ratio} = (\text{Impact} - E(Y))/\sigma(\text{Impact})$$

- Account for small sample bias and non-Gaussian input distributions

$$p\text{-value} = \text{Percentile of } E(Y) \text{ in the distribution of Impact}$$
Stress Testing by Polymodels

Model Selection
- For each subset of indices $I = (i_1, \ldots, i_q)$, merge models as above
- Compute the Information Ratio $= \text{Merged Impact} / \text{Uncertainty}$
- Find the subset $I$ with the highest Information Ratio

Stepwise Regression
- Find the dominant factor™ $i_1$ with highest Information Ratio
- Take this factor as given. Find the second factor $i_2$ such as, jointly with $i_1$, the Information Ratio is maximum, provided it exceeds that of $i_1$ alone.
- Repeat until the Information Ratio cannot be increased
- Try to remove factors while increasing the Information Ratio
- Stop when it is not possible to add, nor remove, factors
Stress Testing by Polymodels

> Given a scenario $x = (x_1, \ldots, x_n)$ we estimate $\Gamma(x) = (\text{cov}_x(X_i, X_j))$ by LOESS regression, using weights $w_x(t)$ that depend on the proximity of the sample data to the scenario.

> **LOESS = LOcally wEighted Scatter plot Smoothing**

> $Y$ is estimated against each $X_i$ by LOESS regression: $Y = \beta_i(x)(X_i - x_i) + \alpha_i(x) + Z_i$

> Let $C(x) = (\beta_i(x)\gamma_i(x))_{i=1..n}$ and $\Lambda(x) = \Gamma(x)^{-1}C(x)$

> Assuming Gaussian inputs, $\beta_i$ estimate is Fischer, $\alpha_i$ is Gaussian, $\Gamma$ is Wishart and $\Lambda$ is inverse Wishart.

> The efficient number of points is $n_{\text{eff}} = \left(\sum \omega_i(x)\right)^2 / \sum \omega_i(x)^2$

> For a given subset of indices $I \subset \{1, \ldots, n\}$, we can make a joint estimate, also by LOESS regression: $Y = \sum_{i \in I} \lambda_i^I(x) (X_i - x_i) + \alpha_i^I + Z_i$

> The estimate is $\hat{y}_i(x) = a_i$ and the variance is: $\text{Var}(Y_i) = C_i^I(x)\Gamma_i(x)C(x) + \text{Var} Z_i$

> The estimation error distribution is: $\text{var} \hat{y}_i(x) = \text{var} Y_i + \text{var} \left[\sum_{i \in I} \lambda_i^I(x)(x_i - \text{E}(X_i))\right]$

> The information ratio is: $J_I(x) = \frac{\hat{y}_i(x)}{\sqrt{\text{var} \hat{y}_i(x)}}$
Multi-Linear Factor Analysis

\[ y = 0.3684x^2 + 0.4048x - 0.0038 \]

\[ R^2 = 0.2925 \]

Fung-Hsieh 7 factors model: misses large negative events

HFR funds returns in Sep 08, Risk analysis as of Aug 08
Nonlinear Dominant Factor™ Analysis

HFR funds Sep 08, Risk analysis as of Aug 08

174 Riskdata factors, adaptive p-value computation, “most relevant” factor

\[ y = 0.8846x - 0.011 \]

\[ R^2 = 0.8121 \]

Pred Sep 08
Diag
Linear (Pred Sep 08)
S&P500 vs. Economic Indicators

Factors: Growth, Inflation (CPI, PPI), Interest Rates, Currency basket, Unemployment, Real Estate Index...
Portfolio Construction by Polymodels

- Assets $Y_1, \ldots, Y_m$
- Risk Factors $X_1, \ldots, X_n$
- Models $Y_j = \varphi_{ij}(X_i) + \varepsilon_{ij}$
- Portfolio $\sum_{j=1}^m w_j Y_j = \sum_{j=1}^m w_j \varphi_{ij}(X_i) + \sum_{j=1}^m w_j \varepsilon_{ij}$

- Portfolio Dominant Factors™: What is the p-value of the portfolio model?
  - The “explained” part is simple, but we need the residual distribution
  - If we know the covariances $<Y_j, Y_k>$ as well as those of nonlinear transforms of the $X_i$’s, the we get the variance of the portfolio, as well as that of the explained part, and finally an estimate of the p-value
  - Instead of covariances of time series, use those of a polymodel sampling: $\varphi_{ij}(x_{ik})$
  - Bayesian techniques face the same difficulty of covariance estimation
Extreme Risk Budgeting

- Budget for the next crisis to secure long-term returns
- During extreme market conditions, monitoring credit and liquidity risk ↔ Hidden Market Risks
- Polymodels and Stress VaR unveil “hidden risks” by including long-term market history into risk analysis, focusing on extremes
- Measuring “hidden” market risk means integrating gamma, long-term factor risk and return smoothing
- Monitoring “hidden” market risk budget implies shifting from static allocations to stable risk budgets per factors, thereby reflecting ALM constraints & long-term views
- These parameters help discriminate between “lucky” managers (who generate returns based on hidden risks), and real talented ones!
An additional risk premium can be extracted by relaxing “normal” risk constraints. Extreme risk budgeted portfolios (FOFIX) outperform Markovitz under business-as-usual conditions – despite exhibiting a lower Sharpe ratio, and simply because they have higher volatility. In troubled times, Markovitz collapses while Dominant Factor™ analysis holds up well.

Pay for “Hedging Costs”, Relaxing “Business as Usual” Risk Constraints
Long-Term Alpha

Poly-models are a “translator” between factors distribution and fund returns distribution

Funds lack long history, but factors went through many crises

Estimate the long-term distribution $f_i(x)dx$ of each selected factor $X_i$
- Estimation should span a full economic cycle (e.g. 40 years) in order to handle all phases

Deduce the long-term distribution of fund returns according to $X_i$

$$g_i(y)dy = f_i(\varphi_i^{-1}(y))/\varphi_i'(\varphi_i^{-1}(y))dy$$

Merge distributions $g_i(y)dy$ according to their relevance into $g(y)dy$

Long-Term Alpha = $E_g(Y)$

Long-Term Ratio = Long-Term Alpha / Stress VaR
Long-Term Alpha & Fund Ranking

Long-Term Alpha converts optionality into expected returns
- Gamma $\leftrightarrow$ Theta

If correctly estimated, it is the same across all factors
- Powerful and robust predictor of long-term returns

Back-test:
- Compute LT-Alpha and StressVaR of each fund at times $t_1, ..., t_n$
- Sort funds by LT-Alpha bucket (or LT-Sharpe bucket)
- Allocate by *extreme risk parity*: weight $= k / \text{StressVaR}$
- Tested over Jan 2005 – Oct 2010 on 500 hedge funds (HFR)
  - Poly-model computed on 36 months, updated every 6 months
  - 4 months delay before investment
  - Comparison with 36M performance and volatility
Sharpe Ratio Selection Iso-Volatility Allocation

- Sharpe > 1
- 0 < Sharpe < 1
- Sharpe < 0
- All

DATA
- Mar'05
- Jun'05
- Sep'05
- Dec'05
- Mar'06
- Jun'06
- Sep'06
- Dec'06
- Mar'07
- Jun'07
- Sep'07
- Dec'07
- Mar'08
- Jun'08
- Sep'08
- Dec'08
- Mar'09
- Jun'09
- Sep'09
- Dec'09
- Mar'10
- Jun'10
- Sep'10

Dominant Factor™ is a protected Trademark of DataCore Innovations, LLC (2015) and had been used with the permission of DataCore Innovations, LLC.
LT-Sharpe Selection Iso-StressVaR Allocation

- LT-Sharpe > 1
- 0 < LT-Sharpe < 1
- LT-Sharpe < 0
- All

DOUADY - NONLINEAR FACTOR ANALYSIS BY POLYMODELS

Dominant Factor™ is a protected Trademark of DataCore Innovations, LLC (2015) and had been used with the permission of DataCore Innovations, LLC
## Fund Ranking

<table>
<thead>
<tr>
<th></th>
<th>E</th>
<th>D</th>
<th>C</th>
<th>B</th>
<th>A</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Med</td>
<td>E</td>
<td>D</td>
<td>C</td>
<td>B</td>
<td>A</td>
</tr>
<tr>
<td>Good</td>
<td>E</td>
<td>D</td>
<td>C</td>
<td>B</td>
<td>A</td>
</tr>
<tr>
<td>Low</td>
<td>E</td>
<td>D</td>
<td>C</td>
<td>B</td>
<td>A</td>
</tr>
<tr>
<td>Risk ↑</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alpha →</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Risk ↑ Alpha →: Bad, Poor, Med, Good, High
LT-Alpha Selection, Iso-Stress VaR Allocation
Portfolio of ETFs

Datasave
Datagrow
Dynamic Leverage
Equity America USA
SP500
Draw Down Datasave
Draw Down Datagrow
Draw Down Dynamic Leverage
Draw Down Equity USA
One of the Most Powerful Buy-side Quant Tool

Dominant Factor™ as a Risk tool
- Stress VaR
- Single Factor Stress Testing with Single factor model
- Full Scenario Testing with Max Information Ratio

Dominant Factor™ as a Replication tool
- Multi-model selection with Max Information Ratio

Dominant Factor™ as a Selection Tool
- Long-term Alpha
- Long-term Ratio

Dominant Factor™ as Allocation Tool
- Extreme Risk-parity
- Extreme Risk Budgeting