DYNAMIC ANALYSIS OF HEDGE FUNDS

Michael Markov
Markov Processes International
428 Springfield Ave., Summit,
NJ 07901, USA
michael.markov@markovprocesses.com

Ilya Muchnik
Rutgers University
P.O. Box 8018, Piscataway,
NJ 08854, USA
muchnik@dimacs.rutgers.edu

Vadim Mottl, Olga Krasotkina
Tula State University
Lenin Ave. 92, Tula, 300600, Russia
vmottl@yandex.ru

ABSTRACT

In this paper, we review one of the most effective financial multi-factor models, called the Returns Based Style Analysis (RBSA), from the standpoint of its performance in detecting dynamic factor exposures of hedge funds using only fund performance data. We analyze the shortcomings of these models in capturing changes in a dynamic portfolio structure and lay the groundwork for a new approach, which we call Dynamic Style Analysis (DSA). The problem is treated as that of estimating a time-varying regression model of the observed time series with the inevitable necessity to choose the appropriate level of model volatility, ranging from the full stationarity of instant models to their absolute independence of each other. We further propose an efficient method of model estimation and introduce a novel measure of the validity Predicted R^2 that is used to select the model parameters. Using both model and real hedge fund returns we illustrate the advantages of the proposed technique in analysis of hedge funds.

KEYWORDS

Style analysis of investment portfolios, hedge funds, time-varying regression, leave-one-out principle, Kalman filter.

1 Introduction

Hedge fund industry has grown rapidly over the past decade to almost $1 trillion in assets and over 8,000 funds. At the same time, the amount of information on hedge funds available to investors is negligible as compared to traditional investment products such as mutual funds. In most cases, the only available information on a hedge fund is a time series of monthly returns and a vague description of the strategy. Returns are then analyzed using a variety of multi-factor models in order to detect the sensitivity of the hedge fund strategy to various risk fac*tors (factor exposures*) as well as to explain the fund's performance in the past.

One of the most effective and practical multi-factor models for analyses of investment portfolios, called the Returns-Based Style Analysis (RBSA), was suggested by Sharpe [1,2]. In the RBSA model, the periodic return of a portfolio is approximately represented by a *constrained linear regression* of a relatively small number of single factors whose role is played by periodic returns of generic market indices each of which represents a certain investment style or sector (market capitalization, quality, duration, region, etc.).

In order to account for allocation changes in active portfolios, Sharpe used a moving window of some preset length [2], assuming that the structure of the portfolio is constant inside the window.

This work is supported by the Russian Foundation for Basic Research, Grant 06-01-00412.

Fung and Hsieh [3] applied RBSA to hedge funds where the method was reduced to unconstrained linear regression to account for shorting and leverage typical for hedge fund strategies. They note a significant loss of explanatory power of RBSA as compared to traditional investment products such as mutual funds (0.25 and 0.75 median R^2 respectively). They conclude that such a low R^2 is due to dynamic nature of hedge fund strategies and introduce generic indices designed to capture hedge fund dynamics, thus increasing median R^2 to 0.4 using the same static regression approach.

In [4], Fung and Hsieh further explore the issue of non-stationarity of RBSA and introduce a method to detect structural breakpoints in factor exposures to improve the R^2, but otherwise the allocations remain constant within the estimation window.

As a generalization of the static RBSA model, we propose here a dynamic model in which a portfolio weights are considered as changing in time. The proposed approach, which we call Dynamic Style Analysis (DSA), consists in estimating a time-varying regression model of the observed time series of periodic returns of the portfolio and generic market indices. Time-varying regression has been subject of intensive study in statistical and econometric literature during, at least, the recent fifteen years [5,6]. In this paper, we consider the problem of estimating a time-varying regression model of a portfolio in its inevitable connection with the necessity to choose the appropriate level of volatility of results, ranging from the full stationarity of instant regression models to their absolute independence of each other. For choosing the volatility level, we use the leave-one-out principle widely adopted in Machine Learning [7,8].

We illustrate the proposed approach to portfolio analysis by applying it to both model and real hedge fund strategies.

2 Sharpe's Returns Based Style Analysis and its limitations

In Sharpe's model, the periodic return on a portfolio $r^{(p)}$ is being approximated by the return on portfolio of assets indices $r^{(i)}$ with weights $(\beta^{(1)},...,\beta^{(n)})$ equal to fractions invested in each asset at the beginning of the period under the assumption that the entire budget is fully spent on the investment:

$$r^{(p)} \cong \alpha + \sum_{i=1}^{n} \beta^{(i)} r^{(i)} . \quad (1)$$

In [2], Sharpe used this model to analyze performance of a group of US mutual funds and determined that a significant portion of a fund return can be explained by small number of assets. In order to estimate parameters of the model (1), Sharpe would take monthly returns on both the portfolio $\{r_t^{(p)}\}$ and asset indexes $\{r_t^{(i)}\}$ for consecutive months

$t = 1, 2, 3, ...$ and solve the following constrained quadratic optimization problem:

$$\begin{cases} (\hat{\alpha}, \hat{\beta}^{(1)}, ..., \hat{\beta}^{(n)}) : \sum_{t=1}^{N} \left(r_t^{(p)} - \alpha - \sum_{i=1}^{n} \beta^{(i)} r_t^{(i)} \right)^2 \to \min, \\ \beta^{(i)} \geq 0, \; \sum_{i=1}^{n} \beta^{(i)} = 1. \end{cases} \quad (2)$$

The resulting coefficients $(\hat{\beta}^{(1)}, ..., \hat{\beta}^{(n)})$ help to identify the major factors determining portfolio performance.

Further, recognizing that portfolio structure changes over time, Sharpe used a series of optimizations in moving windows of a smaller length K to determine the dynamics of portfolio factor exposures:

$$(\hat{\alpha}_t, \hat{\beta}_t^{(1)}, ..., \hat{\beta}_t^{(n)}) = \arg\min_{\alpha, \beta^{(i)}} \sum_{k=0}^{K-1} \left(r_{t-k}^{(p)} - \alpha - \sum_{i=1}^{n} \beta^{(i)} r_{t-k}^{(i)} \right)^2. \quad (3)$$

The model (2) became commonly adopted in the modern Finance under the name of Returns Based Style Analysis (RBSA). The main appeal of this method for practitioners is that it is based solely on analysis of portfolio returns and does not require any other, very often proprietary, information about the portfolio composition.

The two major factors contributing to such wide acceptance of RBSA are its ease of interpretation and stability of results. It is worth noting that both factors are the direct result of the presence of non-negativity constraints in (2). These constraints, being the major innovation in RBSA, provide important prior information about the analyzed portfolio, i.e., the fact that most of investment portfolios such as mutual funds don't take short (negative) positions.

Since its introduction in 1992, Sharpe's model (1) has been criticized for its inability to capture an active portfolio's dynamics. Thus, because portfolio structure is assumed constant within estimation window, moving window technique (3) appears to be inadequate to capture rapid changes in portfolio structure.

In addition, model (1) loses most of its advantages when it is applied to analysis of portfolios that are allowed to take short (negative) positions. In such cases, non-negativity constraints $\beta^{(i)} \geq 0$ have to be dropped from (2), and the problem is reduced to a simple linear regression. In most such cases, due to multicolinearity effect, moving window method (3) produces highly unstable, meaningless results.

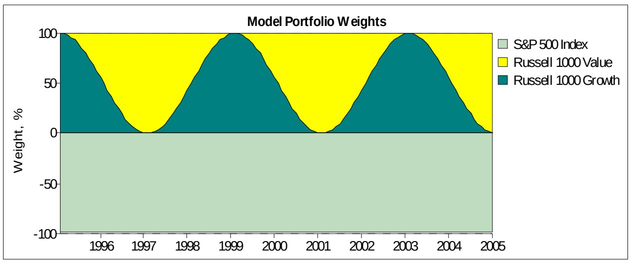
The two limitations above make RBSA inapplicable for analysis of hedge funds because, unlike traditional investment vehicles such as mutual funds, hedge funds are extremely dynamic and take significant short positions.

Most attempts to overcome these shortcomings of RBSA are limited to introduction of additional indices into the static model (1) to capture the specifics of a generic hedge fund strategy [3]. None of the methods available to date represent true dynamic model and, therefore, their explanatory power remains low.

3 Limitations of RBSA: Dynamic model of a hedge fund

We will illustrate the shortcomings of RBSA using a simple model of an equity long-short hedge fund. The long position of such a fund is created using Russell 1000 Value and Growth indices with weights following a sine-wave pattern as shown in Figure 1.

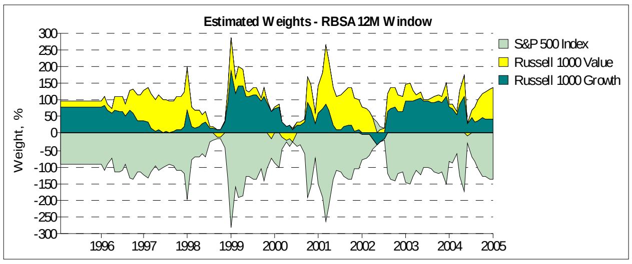
The fund is invested 100% in the Russell 1000 Growth as of Jan-96 and then shifts assets into Russell 1000 Value with a relatively low 50% annual turnover. At any point of time the sum of both index weights is equal to 100%. We then create a long-short model portfolio (MP) by 100% hedging the long portfolio with S&P 500 Index, i.e., effectively subtracting the index returns from the long portfolio return.



Created with MPI Stylus™

Figure 1. The long-short model portfolio.

Next, we perform rolling 12-month window Sharpe's RBSA (3) on the composite monthly return time series of the portfolio MP using, as regressors, the same three monthly indices that were used in its construction. The results are presented in Figure 2 where estimated allocations are stacked along the Y-axis.



Created with MPI Stylus™

Figure 2. Estimated model portfolio, 12-month trailing window.

The results don't materially change when we vary the window size. In Figure 3 we show results of rolling 24-month window.

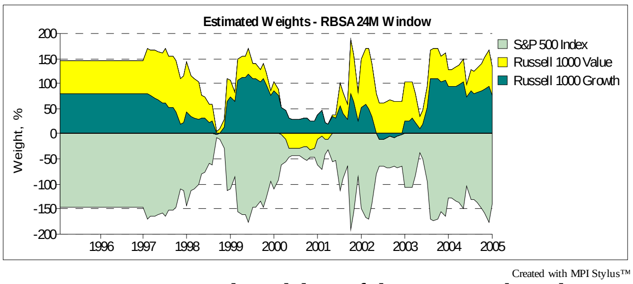


Figure 3. Estimated model portfolio, 24 month trailing window.

Although the turnover of the model portfolio is very low and the number of assets is very small, RBSA fails to adequately identify the model. Since no noise was added to model portfolio returns, this provides clear indication that such poor model identification is the result of the window-based approach and multicolinearity as noted in Section 2, rather than noise in data as it is usually assumed. In Table 1 below we present the correlation matrix of assets used in construction of portfolio MP computed over the same 10-year period using monthly returns. The numbers in brackets represent the range of correlations computed over rolling 24-month windows.

The dynamic model introduced further in this paper eliminates shortcomings of Sharpe's RBSA and makes it applicable to long-short strategies and hedge funds.

Table 1. The correlation matrix of assets constituting the model portfolio.

	R1000G	*R1000V*	*S&P500*
R1000G	1.00		
R1000V	0.71 (0.27;0.95)	1.00	
S&P500	0.94 (0.88;0.99)	0.90 (0.67;0.99)	1.00

4 Dynamic Style Analysis (DSA)

In contrast to Sharpe's static model (1), we propose here a model in which factor exposures of the portfolio are changing in time. Let $t = 1, 2, ..., N$ be a sequence of holding periods, for instance, days, weeks, months, quarters or years, and

$$B = (\boldsymbol{\beta}_t, t = 1, ..., N), \quad \boldsymbol{\beta}_t = (\beta_t^{(0)}, \beta_t^{(1)}, ..., \beta_t^{(n)}), \quad \sum_{i=0}^{n} \beta_t^{(i)} = 1,$$

be the respective sequence of the portfolio's exposures at the end of each period. The notation $\beta_t^{(0)}$ is reserved here for a short-term instrument, such as bank deposit in an interest bearing account, often referred to as risk-free asset.

For simplicity, we will express the model in terms of excess returns on the portfolio $(r_t^{(p)} - r_t^{(0)})$ and assets $(r_t^{(i)} - r_t^{(0)})$ with respect to the return on the risk-free asset $r_t^{(0)}$. Such an equivalent notation effectively eliminates the need for the budget constraint $\beta_t^{(0)} + \sum_{i=1}^{n} \beta_t^{(i)} = 1$ in (1). The new dynamic model of periodic portfolio returns can be written as follows:

$$y_t = (r_t^{(p)} - r_t^{(0)}) = \sum_{i=1}^{n} \beta_t^{(i)} (r_t^{(i)} - r_t^{(0)}) = \sum_{i=1}^{n} \beta_t^{(i)} x_t^{(i)} + e_t = \boldsymbol{\beta}_t^T \mathbf{x}_t + e_t \quad .(4)$$

Here $y_t = (r_t^{(p)} - r_t^{(0)})$ are known excess returns of the portfolio for each period t, and $\mathbf{x}_t = \left[(r_t^{(i)} - r_t^{(0)}), i = 1, ..., n \right] \in \mathbb{R}^n$ are known vectors of observed excess returns of assets for these periods, whereas $\boldsymbol{\beta}_t = (\beta_t^{(1)}, ..., \beta_t^{(n)}) \in \mathbb{R}^n$ are vectors of time-varying fractional asset weights to be estimated.

The key element of the proposed *Dynamic Style Analysis* (DSA) is treating fractional asset weights as a hidden process assumed *a priori* to possess the Markov property:

$$\boldsymbol{\beta}_t = \mathbf{V}_t \boldsymbol{\beta}_{t-1} + \boldsymbol{\xi}_t, \qquad (5)$$

where matrices $\mathbf{V}_t$ determine the assumed hidden dynamics of the portfolio structure, and $\boldsymbol{\xi}_t$ is the vector white noise, nonstationary in the general case.

The equation (5) determines the state-space model of a dynamic system, while (4) plays the role of its observation model. In these terms, the DSA problem can be described as estimating the time-varying state of the respective dynamic system $B(Y, X) = \left(\boldsymbol{\beta}_t (Y, X), t = 1, ..., N \right)$ from observations $(Y, X) = \left[(y_t, \mathbf{x}_t) = \left((r_t^{(p)} - r_t^{(0)}), (r_t^{(i)} - r_t^{(0)}), i = 1, ..., n \right), t = 1, ..., N \right]$.

For estimating time-varying models of kind (4)-(5), we use the *Flexible Least Squares* approach (FLS) first introduced in [5]. As applied to the DSA problem, the FLS criterion has the form

$$\begin{cases} \hat{B}(Y, X, \lambda) = \arg\min J(\boldsymbol{\beta}_t, t = 1, ..., N \mid Y, X), \\ J(\boldsymbol{\beta}_t, t = 1, ..., N \mid Y, X) = \\ \sum_{t=1}^{N} (y_t - \boldsymbol{\beta}_t^T \mathbf{x}_t)^2 + \lambda \sum_{t=2}^{N} (\boldsymbol{\beta}_t - \mathbf{V}_t \boldsymbol{\beta}_{t-1})^T \mathbf{U}_t (\boldsymbol{\beta}_t - \mathbf{V}_t \boldsymbol{\beta}_{t-1}). \end{cases} \qquad (6)$$

The assumed covariance matrices $\mathbf{Q}_t$ of white noise $\boldsymbol{\xi}_t$ in (5) occur here in the inversed form $\mathbf{U}_t = \mathbf{Q}_t^{-1}$. We shall additionally assume matrices $\mathbf{V}_t$ to be non-degenerate, in this case, they will determine also the reversed dynamics of the time-varying regression coefficients.

The positive parameter λ in (6) is responsible for the noise ratio in (4)-(5), i.e. for the level of smoothness of regression coefficients. Thus, the smoothness parameter λ balances the two conflicting requirements: to provide close approximation of portfolio returns and, at the same time, to control the smoothness of asset weights $\boldsymbol{\beta}_t$ over time.

5 The cross validation principle of estimating the smoothness parameter

The FLS criterion (6) depends on a number of parameters. Matrices $\mathbf{V}_t$ and $\mathbf{U}_t$ of the transition model (5) can be defined a priori, depending on the model of hidden dynamics of time-varying regression coefficients and the "style" of smoothing of their estimates. For example, matrix $\mathbf{V}_t$ can be defined as the unity matrix thus requiring simple smoothness of estimates. Alternatively, we could allow for nonsmoothness of asset weights by incorporating market-driven changes of weights into transition model (5) as follows:

$$\beta_t^{(i)} = \frac{1+x_{t-1}^{(i)}}{\sum_{k=1}^{n}\beta_{t-1}^{(k)}x_{t-1}^{(i)}}\beta_{t-1}^{(k)} + \xi_t^{(i)},\ i=1,...,n,\ t=2,...,N .$$

As to the coefficient λ, it is extremely problematic to preset its value a priori. At the same time, there is a deep specificity in estimating it from the observed time series.

If the volatility parameter is given a certain value λ, the FLS estimate of time-varying regression coefficients (6) will be a function of it $\hat{B}(Y,X,\lambda) = \left(\hat{\boldsymbol{\beta}}_t(Y,X,\lambda), t=1,...,N\right)$. It is impossible to find an "appropriate" value of λ by attempting to additionally minimize the residual squares sum in (6) $\sum_{t=1}^{N}\left(y_t - \hat{\boldsymbol{\beta}}_{t-1}^T(Y,X,\lambda)\mathbf{x}_t\right)^2 \to \min_\lambda$.

Indeed, when $\lambda \to \infty$, the second sum in (6) totally prevails over the first sum, the values of the hidden process become functionally related to each other $\hat{\boldsymbol{\beta}}_t(Y,X,\lambda) = \mathbf{V}_t\hat{\boldsymbol{\beta}}_{t-1}(Y,X,\lambda)$, and the model is reduced to a static regression. If, on the contrary, $\lambda \to 0$, the instantaneous values are getting a priori independent, each estimate $\hat{\boldsymbol{\beta}}_{t-1}(Y,X,\lambda)$ is determined practically by only one current element of the time series $(y_t, \mathbf{x}_t)$, and the model will be "extremely" time-varying.

Actually, the sought-for sequence of time-varying regression coefficients $B(Y,X,\lambda) = \left(\boldsymbol{\beta}_t(Y,X,\lambda), t=1,...,N\right)$ is a model of the observed time series $(Y,X) = \left[(y_t, \mathbf{x}_t) = \left((r_t^{(p)} - r_t^{(0)}), (r_t^{(i)} - r_t^{(0)}), i=1,...,n\right), t=1,...,N\right]$, and the choice of λ is the choice of a class of models which would be most adequate to the data. A commonly used measure of regression model fit is its coefficient of determination R^2. This coefficient was computed by Sharpe in his work [2] as the proportion of the portfolio volatility explained by systematic exposures using the moving window technique (3). In terms of the FLS criterion (6), the coefficient of determination will be expressed as the ratio

$$R^2 = \frac{\sum_{t=1}^{N}(y_t)^2 - \sum_{t=1}^{N}\left(y_t - \hat{\boldsymbol{\beta}}_t^T(\lambda)\mathbf{x}_t\right)^2}{\sum_{t=1}^{N}(y_t)^2} = 1 - \frac{\sum_{t=1}^{N}\left(y_t - \hat{\boldsymbol{\beta}}_t^T(\lambda)\mathbf{x}_t\right)^2}{\sum_{t=1}^{N}(y_t)^2}. \quad (7)$$

However, by decreasing λ, it is easy to drive R^2 up to 100% and, at the same time, obtain highly volatile, meaningless estimates of fractional asset weights $\hat{\boldsymbol{\beta}}_t(Y,X,\lambda)$.

The major reason for such inadequacy of the R^2 statistic is that it uses the same data set for both estimation and verification of the model. The *Cross Validation* method suggested by Allen [9] under the name of *Prediction Error Sum of Squares* (PRESS) is aimed at overcoming this obstacle. According to this method, an observation is removed from the sample, the model is evaluated on the remaining observations, and the prediction error is calculated on the removed observation. This procedure is then repeated for each observation in the sample, and the sum of squared errors is computed. The Cross Validation principle is widely adopted in data analysis [10,11], including pattern recognition, where it is known under the name of the "leave-one-out" procedure [8,12].

As applied to verification of the accuracy of the FLS model, the essence of the Cross Validation principle can be explained as the idea to assess the adequacy of the given model by estimating the variance of the residual noise $D(e)$ in (4) and comparing it with the full variance of the goal variable $D(y) = (1/N)\sum_{t=1}^{N}(y_t)^2$. But, when computing the error at a time moment t, it is incorrect to use the estimate $\hat{\boldsymbol{\beta}}_t$ obtained by minimizing the criterion (6) with participation of the observation at this time moment $(y_t, \mathbf{x}_t)$. The CV principle leads to the following procedure that provides a correct estimate of the observation noise variance.

In the full time series $\left((y_1,\mathbf{x}_1),...,(y_N,\mathbf{x}_N)\right)$, single elements $t=1,...,N$ are skipped one by one $\left((y_1,\mathbf{x}_1),..., (y_{t-1},\mathbf{x}_{t-1}), (y_{t+1},\mathbf{x}_{t+1}),...,(y_N,\mathbf{x}_N)\right)$, each time by replacing the sum $\sum_{t=1}^{N}\left[y_t - \left(\boldsymbol{\beta}_t(\lambda)\right)^T\mathbf{x}_t\right]^2$ in (6) by the truncated sum $\sum_{s=1,s\neq t}^{N}\left[y_s - \left(\boldsymbol{\beta}_s(\lambda)\right)^T\mathbf{x}_s\right]^2$, and the optimal vector sequences $(\hat{\boldsymbol{\beta}}_1^{(t)},...,\hat{\boldsymbol{\beta}}_N^{(t)})$ are found, where the upper index (t) means that the observation $(y_t,\mathbf{x}_t)$ was omitted when computing the respective estimate. For each t, the instantaneous squared prediction error is calculated using the respective single estimate $\left[y_t - \left(\hat{\boldsymbol{\beta}}_t^{(t)}(\lambda)\right)^T\mathbf{x}_t\right]^2$. The cross-validation estimate of the noise variance is found as the average over all the local squared prediction errors

$$\hat{D}_{CV}(e\,|\,\lambda) = \frac{1}{N}\sum_{t=1}^{N}\left[y_t - \left(\hat{\boldsymbol{\beta}}_t^{(t)}(\lambda)\right)^T\mathbf{x}_t\right]^2 \quad (8)$$

The less $\hat{D}_{CV}(e\,|\,\lambda)$, the more adequate is the model with the given value of the smoothness parameter λ to the observed time series $\left((y_1,\mathbf{x}_1),...,(y_N,\mathbf{x}_N)\right)$.

The cross-validation estimate of the residual noise variance $\hat{D}_{CV}(e\,|\,\lambda)$ can be further scaled to make it comparable across different analyzed portfolios. We suggest the cross-validation statistic

$$PR^2(\lambda) = \frac{D(y) - \hat{D}_{CV}(e\,|\,\lambda)}{D(y)} = 1 - \frac{\hat{D}_{CV}(e\,|\,\lambda)}{D(y)}. \quad (9)$$

which we call *Predicted R-squared*. Note that it is computed similarly to the regression R-squared statistic (7).

We suggest a method of determining optimal model parameters that consists in processing the given time series $\left((y_1,\mathbf{x}_1),...,(y_N,\mathbf{x}_N)\right)$ several times with different tentative values of λ. Each time, the model adequacy is assessed by the averaged squared prediction error (8) estimated by the cross validation procedure. The value λ^* that yields the maximum value of the cross-validation statistic (9) is to be taken as the smoothing parameter recommended for the given time series:

$$\lambda^* = \arg\max_\lambda PR^2(\lambda). \quad (10)$$

It should be noted that the selection of model parameters through minimizing the prediction error makes this method a version of the James-Stein estimator producing the smallest prediction error [10].

6 Kalman filter and smoother for minimization of flexible least squares and cross validation

The FLS criterion (6) is a quadratic function, and its minimization leads to a system of linear equations. At the same time, it belongs to the class of so-called pair-wise separable optimization problems [13], in which the objective function is the sum of functions each depending on not more than two vector variables, in this case $\boldsymbol{\beta}_{t-1}$ and $\boldsymbol{\beta}_t$ associated with immediately successive time moments. As a result, the matrix of the linear equation system relative to variables $\boldsymbol{\beta}_1,...,\boldsymbol{\beta}_N$ will have block-threediagonal structure, which allows for easily solving it by the double-sweep method, which is a quadratic version of the much more general dynamic programming method [13]. These equivalent algorithms are, in their turn, equivalent to the Kalman filter and smoother [14].

First, the Kalman filter runs along the signal

$$\boldsymbol{\beta}_{1|1}=\left(y_1/\mathbf{x}_1^T\mathbf{x}_1\right)\mathbf{x}_1,\ \mathbf{Q}_{1|1}=\mathbf{x}_1\mathbf{x}_1^T \text{ at } t=1,$$

$$\boldsymbol{\beta}_{t|t}=\mathbf{V}_t\boldsymbol{\beta}_{t-1|t-1}+\mathbf{Q}_{t|t}^{-1}\mathbf{x}_t(y_t-\mathbf{x}_t^T\mathbf{V}_t\boldsymbol{\beta}_{t-1|t-1}),\ t=2,...,N, \quad (11)$$

$$\mathbf{Q}_{t|t}=\mathbf{x}_t\mathbf{x}_t^T+\mathbf{U}_t\mathbf{V}_t\left(\mathbf{V}_t^T\mathbf{U}_t\mathbf{V}_t+(1/\lambda)\mathbf{Q}_{t-1|t-1}\right)^{-1}\mathbf{Q}_{t-1|t-1}\mathbf{V}_t^{-1}= \mathbf{x}_t\mathbf{x}_t^T+\left(\mathbf{V}_t\mathbf{Q}_{t-1|t-1}\mathbf{V}_t^{-1}+(1/\lambda)\mathbf{U}_t^{-1}\right). \quad (12)$$

The intermediate vectors $\boldsymbol{\beta}_{t|t}$ and matrices $\mathbf{Q}_{t|t}$ are parameters of the so-called Bellman functions $J_{t|t}(\boldsymbol{\beta}_t)$ being quadratic in this case [14]:

$$J_{t|t}(\boldsymbol{\beta}_t)=\min_{\boldsymbol{\beta}_1,...,\boldsymbol{\beta}_{t-1}}J_t(\boldsymbol{\beta}_1,...,\boldsymbol{\beta}_t)=(\boldsymbol{\beta}_t-\boldsymbol{\beta}_{t|t})^T\mathbf{Q}_{t|t}(\boldsymbol{\beta}_t-\boldsymbol{\beta}_{t|t})+const,$$

$$J_t(\boldsymbol{\beta}_1,...,\boldsymbol{\beta}_t)=\sum_{s=1}^t(y_s-\boldsymbol{\beta}_s^T\mathbf{x}_s)^2+\lambda\sum_{s=2}^t(\boldsymbol{\beta}_s-\mathbf{V}_s\boldsymbol{\beta}_{s-1})^T\mathbf{U}_s(\boldsymbol{\beta}_s-\mathbf{V}_s\boldsymbol{\beta}_{s-1}).$$

Here $J_t(\boldsymbol{\beta}_1,...,\boldsymbol{\beta}_t)$ are partial criteria of the same structure as (6). The minimum points $\boldsymbol{\beta}_{t|t}$ of the Bellman functions yield the so-called filtration estimates of the unknown regression coefficients at current t under the assumption that the time series is observed only up to point t.

Then, the Kalman smoother runs in the backward direction $t=N-1,...,1$:

$$\hat{\boldsymbol{\beta}}_t=\boldsymbol{\beta}_{t|t}+\mathbf{H}_t(\hat{\boldsymbol{\beta}}_{t+1}-\boldsymbol{\beta}_{t|t}), \quad (13)$$

$$\mathbf{H}_t=\left(\mathbf{V}_{t+1}^T(\lambda\mathbf{U}_{t+1})\mathbf{V}_{t+1}+\mathbf{Q}_{t|t}\right)^{-1}\mathbf{V}_{t+1}^T(\lambda\mathbf{U}_{t+1})\mathbf{V}_{t+1}= \left(\mathbf{I}+(1/\lambda)\mathbf{V}_{t+1}^{-1}\mathbf{U}_{t+1}^{-1}(\mathbf{V}_{t+1}^{-1})^T\mathbf{Q}_{t|t}\right)^{-1}. \quad (14)$$

The resulting sequence is just the minimum point of the FLS criterion (6) $\hat{B}(Y,X,\lambda)=\left(\hat{\boldsymbol{\beta}}_t(Y,X,\lambda),\ t=1,...,N\right)$.

To compute the leave-on-out estimate of the noise variance (8), we have to find the estimate of each regression coefficient vector $\hat{\boldsymbol{\beta}}_t(Y^{(t)},X^{(t)},\lambda)$ from the time series $(Y^{(t)},X^{(t)})=\left((y_s,\mathbf{x}_s),\ s=1,...,t-1,t+1,...,N\right)$ where the element $(y_t,\mathbf{x}_t)$ is cut out. This means that, when running the Kalman filter, we have to use at step t the matrix $\mathbf{Q}_{t|t}^{(t)}=\mathbf{Q}_{t|t}-\mathbf{x}_t\mathbf{x}_t^T=\left(\mathbf{V}_t\mathbf{Q}_{t-1|t-1}\mathbf{V}_t^{-1}+(1/\lambda)\mathbf{U}_t^{-1}\right)^{-1}$ instead of $\mathbf{Q}_{t|t}$ (12).

7 The Computational Complexity of DSA

However, straightforward application of the cross validation principle (8)-(10) to determining value of the smoothness parameter implies running the Kalman filtration-smoothing procedure (11)-(14) N times for each removed observation corresponding to time period t, which fact will destroy the linear computational complexity of the algorithm with respect to the length of the time series N. Thus, to analyze $N=120$ monthly returns of a portfolio using $n=10$ economic sectors as variables, it is required to solve a quadratic problem (6) with $Nn=1{,}200$ variables, what can be easily done by the standard Kalman filter-smoother having linear computational complexity with respect to N. But in order to compute the cross-validation statistic (9) corresponding to a single value of the smoothness parameter, $N=120$ such optimizations are required, and computing the CV statistic on a grid of 20 values of this parameter requires solving $20N=2{,}400$ problems (6), i.e. 2,400 runs of the optimization procedure.

To avoid repeated processing of the signal for each tentative value of λ, a technique of incorporating computation of the leave-one-out error (8) into the main dynamic programming procedure is proposed in [14]. It is shown that the estimate $\hat{\boldsymbol{\beta}}_t^{(t)}(\lambda)$ is determined by the expression

$$\hat{\boldsymbol{\beta}}_t^{(t)}(\lambda)=\hat{\boldsymbol{\beta}}_t(Y^{(t)},X^{(t)},\lambda)=(\mathbf{Q}_{t|N}-\mathbf{Q}_t^0)^{-1}(\mathbf{Q}_{t|N}\boldsymbol{\beta}_{t|N}-\mathbf{Q}_t^0\boldsymbol{\beta}_t^0),$$

where matrices $\mathbf{Q}_{t|N}$ are to be additionally computed at the backward run of the Kalman smoother for $t=N-1,...,1$

$$\mathbf{Q}_{t|N}=\left(\mathbf{H}_t\mathbf{V}_{t+1}^{-1}\mathbf{Q}_{t+1|N}^{-1}(\mathbf{V}_{t+1}^{-1})^T\mathbf{H}_t^T+(\mathbf{Q}_{t|t}+\lambda\mathbf{V}_{t+1}^T\mathbf{U}_{t+1}\mathbf{V}_{t+1})^{-1}\right)^{-1},$$

starting with matrix $\mathbf{Q}_{N|N}$ found at the last step of the Kalman filter (12).

8 Testing the DSA Approach: Dynamic model of a hedge fund

We applied the DSA approach (6) to the model portfolio MP developed in Section 3 with the smoothness parameter λ selected in accordance with (10). The result of such analysis is shown in Figure 4.

In order to test sensitivity of the model to noise in data, we added an idiosyncratic white noise to the MP monthly returns in the amount of 20% of the MP volatility[1]. The resulting portfolio returns were analyzed, and the result corresponding to the maximum value of the CV statistic is shown in Figure 5. The result corresponds to the optimal smoothness parameter $\lambda=0.2$ selected to provide the maximum value of the PR^2 statistic (10).

[1] Since the standard deviation of MP monthly returns over 120 months makes $\sigma=1.08$, we applied white noise $N(0,0.22)$.

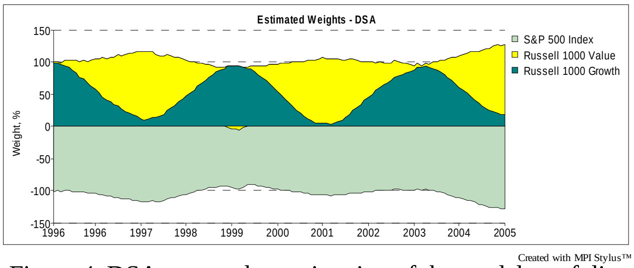


Figure 4. DSA approach – estimation of the model portfolio.

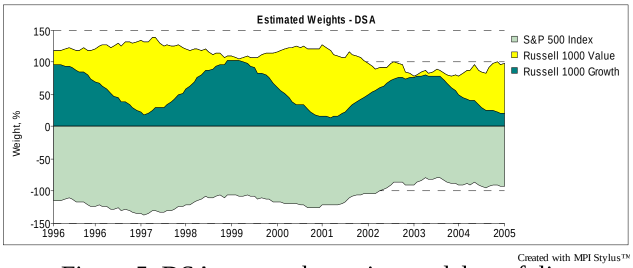


Figure 5. DSA approach – noisy model portfolio.

9 Case Studies

In this section, we present examples of the application of the dynamic multi-factor methodology developed in this paper to real-life hedge funds.

9.1 Laudus Rosenberg Value Long/Short Fund

According to the fund prospectus, the Laudus Rosenberg Value Long/Short mutual fund[1] was using computer models for buying underpriced US stocks and selling stocks short in order to maintain both market and sector neutrality. Such neutrality is very important for investors because it protects their investment in market downturns.

Fund monthly returns are shown in Figure 6. We will compare performance of Sharpe's RBSA and Dynamic Style Analysis (DSA) in determining sensitivity of the fund returns to economic sectors.

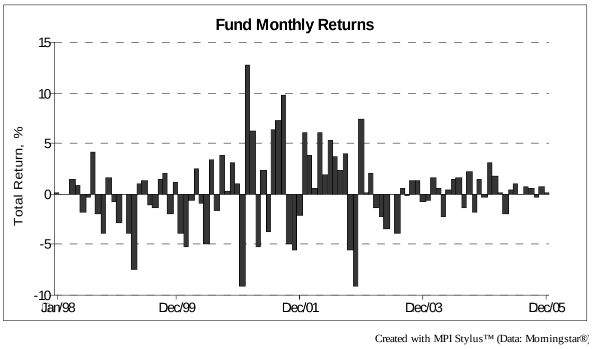


Figure 6. Fund returns.

For our analysis we used 10 indexes provided by Dow Jones Indexes[2]. The result of the analysis using a 36-month window RBSA (3) is presented on the Figure 7.

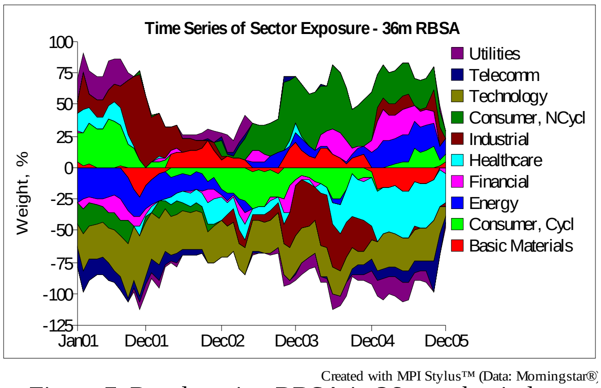


Figure 7. Results using RBSA in 36 month window.

The result is very volatile and unrealistic with $R^2 = 0.60$ (7). Shortening the estimation window produces higher R^2 value but much more volatile results.

We then applied the DSA approach using the same return data and selected the optimal parameter λ in accordance with (10). The resulting sector exposures are presented in Figure 8.

In Figure 9 we present the values of criterion $PR^2(\lambda)$ (10) corresponding to various λ plotted along logarithmic axis and the optimal value of the smoothness parameter λ^*. Even though the weights are much less volatile in Figure 8 than the ones in Figure 7 obtained using 36-month rolling RBSA, the resulting $R^2 = 0.86$ (7) is much greater than 0.6, determined by a trailing window. The corresponding optimal value of *Predicted* R^2 (10) is $PR^2(\lambda^*) = 0.53$.

Therefore, the proposed technique allowed us to achieve a much closer approximation of the fund return pattern with a significantly more realistic pattern of sector exposures.

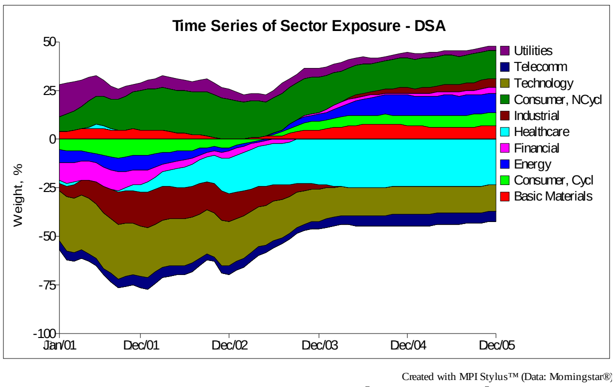
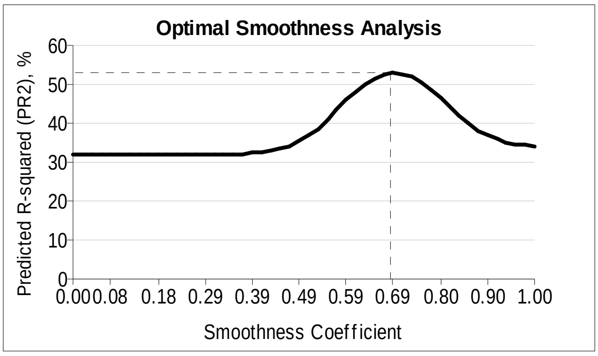


Figure 8. DSA-estimated asset weights.

Therefore, the proposed technique allowed us to achieve a much closer approximation of the fund return pattern with a significantly more realistic pattern of sector exposures.

[1] Laudus Rosenberg Value Long/Short (Ticker: BRMIX) is a mutual fund employing a strategy similar to a long-short hedge fund. Information on this fund is available on finance.yahoo.com and www.morningstar.com

[2] Source: indexes.dowjones.com



Created with MPI Stylus™ (Data: Morningstar®)

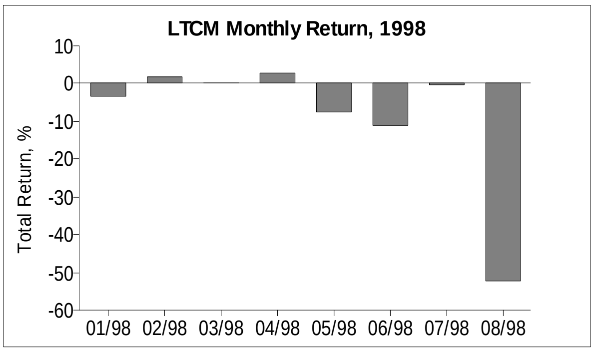
Figure 9. Selection of smoothness parameter.

9.2 Long Term Capital Management (LTCM)

The collapse of this highly-leveraged fund in 1998 is by far the most dramatic hedge fund story to date. At the beginning of 1998, the $5B fund maintained leverage ratio of about 28:1 [15,16,17]. LTCM troubles began in May-June 1998. By the end of August, the fund had lost 52%. By the end of September, 1998, a month after the Russian crisis, the fund lost 92% of its December 1997 assets. Fearing the destabilizing impact of the highly leveraged fund on global financial markets, on September 23rd, the Federal Reserve Bank of New York orchestrated a bailout of the fund by a group of 14 major banks.

The investment strategies of the fund were based on betting on spreads between thousands of closely related securities in various global markets. Such bets are based on the assumption that the securities will eventually converge and the arbitrage position will result in a profit. In fact, the spreads continued to rise which eventually led to collapse of the fund. It took several major investigations, including one commissioned by President Clinton [16], to determine that major losses sustained by LTCM came from bets on global credit spreads.

We will use DSA methodology to determine major factors explaining losses of LTCM in 1998 using the fund's monthly returns. We will also determine the leverage ratio, an important risk factor which, according to published figures [15,16], had increased from 28:1 to 52:1 over the course of 1998. Fund's 1998 monthly returns in January-August shown in Figure 10 were obtained from public sources [17]. Daily or weekly returns, if available, would provide far greater accuracy.



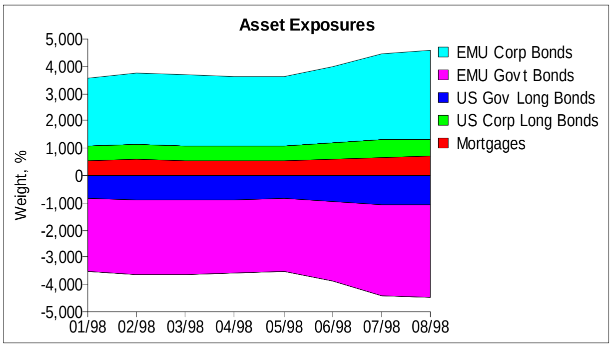
Created with MPI Stylus™

Figure 10. LTCM monthly returns in 1998.

For our analysis we used the following indexes provided by Lehman Brothers and Merrill Lynch: US Corporate and Government Long Bond Indices, European (EMU) Corporate and Government Bond Indices and US Mortgage-Backed Securities Index. The result of the analysis is presented in the Figure 11. Asset exposures $\boldsymbol{\beta}_t$ of the fund for each time period are "stacked" along the Y-axis, with the sum equal to 100%. The negative weights shown below the X-axis correspond to borrowed assets and represent leverage. Evidently, the leverage comes from credits spreads – both US (Corp Index vs. Govt Index) and EMU (Corp Index vs. Govt Index). There's also significant exposure to Mortgages. Our results show that the leverage increased from 35:1 (or 3,500%) to 45:1 (4,500%) during 1998, which is close to the figures published in [16,17].

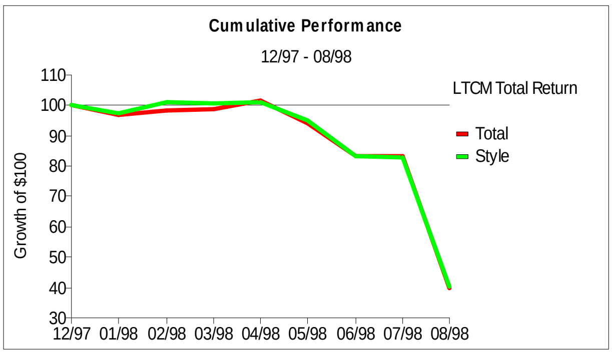
The result corresponds to the optimal smoothness coefficient λ which was selected to provide the maximum value of the Predicted R^2 statistic. The R^2 of this result is 0.99, while Predicted R^2 is 0.98.

The "growth of $100" chart in Figure 12 showing the performance of LTCM in 1998 presents an excellent model fit, where the performance of the fund is very closely approximated by the model. In chart, the "Total" line represents the fund and the "Style" represents the model-based approximation.



Created with MPI Stylus™

Figure 11. LTCM Estimated Asset Weights $\boldsymbol{\beta}_t$.



Created with MPI Stylus™

Figure 12. LTCM performance tracking.

Following Allen's use of leave-one-out PRESS statistic in model selection [9], in Table 2 we use cross-validation static Predicted R^2 (10), to illustrate "optimality" of the result obtained in our analysis of the LTCM. In the table, we show the impact on Predicted R^2 from either removal of each of the model assets or adding new assets. It shows that the full model with 5 assets is preferable, having the highest Predicted R^2 equal to 0.98.

Table 2. Model selection in the LTCM analysis.

Asset Name	Max *Predicted* R^2	
	Asset removed	Asset added
EMU Corp Bonds	0.861	
EMU Govt Bonds	0.857	
US Gov Long Bonds	0	
US Corp Long Bonds	0.839	
Mortgages	0.965	
European Stocks		0.948
US Small Stocks		0.977

We then computed monthly VaR (Value-at-Risk) corresponding to asset exposures in Figure 11 using two years of monthly returns on asset indices employed in the analysis. Depending on parameters of calculation (such as decay factor, distribution assumptions, etc.), the 99% systematic VaR for June-August 1998 is in the 30%-55% range. Therefore, 10-50% monthly losses sustained by the fund during this period should have been expected if proper VaR methodology was used. As it was mentioned above, applying DSA to higher frequency data (daily or weekly) could have produced much more accurate estimates of potential daily losses.

9.3 Replicating a Hedge Fund Index

The purpose of this section is to demonstrate how DSA methodology developed in previous sections could be employed to replicate the performance of a hedge fund strategy index using generic asset indices.

Most hedge fund database vendors publish performance of hedge fund strategy indices – weighted aggregates of funds within groups representing similar investment strategy. Indices representing category averages are readily available from a number of hedge fund database vendors[1]. For our analysis we used monthly returns of the HFR Equity Hedge Index representing Long/Short category, which is one of the most representative. Below, we provide the definition of the category from HFR website:

Equity Hedge investing consists of a core holding of long equities hedged at all times with short sales of stocks and/or stock index options. Some managers maintain a substantial portion of assets within a hedged structure and commonly employ leverage. Where short sales are used, hedged assets may be comprised of an equal dollar value of long and short stock positions. Other variations use short sales unrelated to long holdings and/or puts on the S&P 500 index and put spreads. Conservative funds mitigate market risk by maintaining market exposure from zero to 100 percent. Aggressive funds may magnify market risk by exceeding 100 percent exposure and, in some instances, maintain a short exposure. In addition to equities, some funds may have limited assets invested in other types of securities.

The HFR Equity Hedge index represents an equal-weighted composite of 615 funds within the category. Even though individual hedge funds engage in frequent trading and utilize derivatives, our contention is that in the index, specific risk is diversified and its returns could be explained by a handful of systematic factors. Since most of hedge funds in the category invest in equities, we used the following indices for our analysis: 6 Russell Equity Indices (Top 200 Value/Growth, Midcap Value/Growth, Russell 2000 – Small Cap Value/Growth) as proxies for US Equities, and MSCI EAFE Index as the proxy for international equities and ADRs. We used Merrill Lynch 3-Month TBill index as a proxy for cash. Monthly returns for 7 years July 1999 – June 2006 of both the hedge fund index and generic asset indices were used.

The results of DSA analysis (6) corresponding to the optimal value of parameter λ is shown in Figure 13. The quality of regression fit is very high: $R^2 = 0.98$ and *Predicted* $R^2 = 0.90$.

In Figure 14 we present values of criterion $PR^2(\lambda)$ corresponding to various values of coefficient λ plotted along logarithmic X-axis. Note that the results in Figure 13 were obtained using λ corresponding to the highest $PR^2(\lambda)$. The analysis of exposure levels in Figure 13 present several interesting observations. First, the average leverage level in this category (as measured by the magnitude of short exposures below the X-axis) is relatively small and stable. Note also that market exposure has increased dramatically over 1995-1996 (especially to international equity markets represented by EAFE index), almost to year 2000 levels.

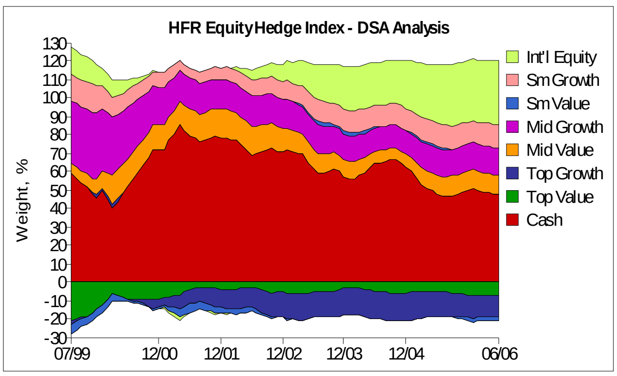
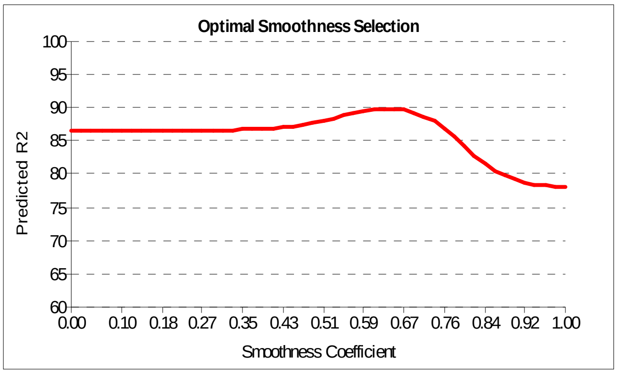


Figure 13. Equity Hedge Index: DSA analysis

[1] Among the most widely used: HFR (Hedge Fund Research) www.hedgefundresearch.com, CSFB/Tremont www.hedgeindex.com, Eurekahedge www.eurekahedge.com, and others.

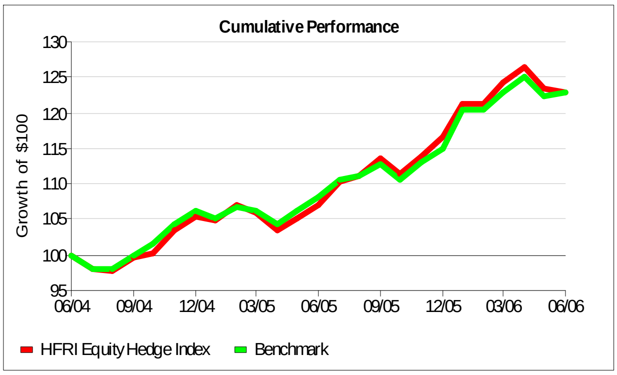


Created with MPI Stylus™

Figure 14. Equity Hedge Index: smoothness selection.

Following described above "in-sample" replication of the Equity Hedge index, we used the same methodology to replicate the index "out-of-sample." We used the first 60 months of data July 1999 through June 2004 to determine allocations to generic indices as of June 2004. We then computed return for the replication portfolio of generic indices for July 2004 using generic index returns for July 2004 and allocations estimated via DSA. We then expanded the estimation interval to include 61 months through July 2004 and estimated replication portfolio return for August 2004. We then repeated the process for each of the remaining months through June 2006.

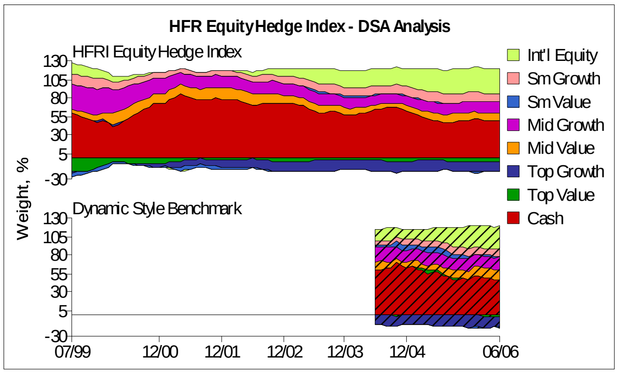
In Figure 15 we show cumulative performance for the HFR Equity Hedge Index and its "out-of-sample" replication ("Benchmark") for two years July 2004 – June 2006. It is clear that index performance has been replicated very closely.



Created with MPI Stylus™

Figure 15. Equity Hedge Index: performance replication

In Figure 16 we compare allocations of both "in-sample" DSA analysis of the Equity Hedge Index (equivalent to the one in Figure 13) and allocations of its "out-of-sample" replication. Note that the latter starts only 60 months after the start of the data sample and the asset weight estimates are more volatile than the former.



Created with MPI Stylus™

Figure 16. Equity Hedge Index: exposures replication.

10 Conclusions

The DSA approach has made it possible to fundamentally improve the existing RBSA methodology currently employed in Finance. The proposed technique that implements existing quadratic optimization algorithms is very practical and can be executed on a personal computer. We provide a framework for truly dynamic analysis of hedge funds, resulting in increased transparency and better due diligence which is of crucial importance for financial institutions today.

However, application of DSA can achieve 100% fit using any explanatory variables by adjusting the smoothness parameter. These variables can be totally unrelated to the analyzed portfolio return. Moreover, even when explanatory variables are selected properly, changing the smoothness parameter can result in very different results. It is therefore required that this parameter is estimated from data, because in most cases analysts don't have information about underlying hedge fund positions and their dynamics.

Thus, there is a need to measure the model adequacy. Clearly, the coefficient of determination R^2 (7) is not suitable. Instead, the leave-one-out cross-validation and *Predicted* R^2 approach solve both above-mentioned issues. We demonstrated this using a model hedge fund portfolio. Aside from its use in the parameter selection above, the *Predicted* R^2 serves a measure of the model validity. Similarly to the PRESS statistic that it is based on, it can be used to select the best set of factors as the one providing the highest value of the *Predicted* R^2.

A modification of the Kalman filter-smoother by incorporating the leave-one-out procedure has allowed us to escape the seemingly unavoidable loss of the linear computational complexity with respect to the length of the time series.

References

[1] W.F. Sharpe. Determining a fund's effective asset mix. *Investment Management Review*, Vol. 2, 1988, No. 6 (December), pp. 59-69.

[2] W.F. Sharpe. Asset allocation: Management style and performance measurement. *The Journal of Portfolio Management*, Winter 1992, pp. 7-19.

[3] W. Fung, D.A. Hsieh. Empirical characteristics of dynamic trading strategies: The case of hedge funds. *Review of Financial Studies*, 1997, V.10, pp. 275–302.

[4] W. Fung, D.A. Hsieh. Hedge fund benchmarks: A risk based approach. *Financial Analysts Journal*, 2004, Vol. 60, p. 65-80.

[5] R. Kalaba, L. Tesfatsion. Time-varying linear regression via flexible least squares. *International Journal on Computers and Mathematics with Applications*, 1989, Vol. 17, pp. 1215-1245.

[6] C. Wells *The Kalman Filter in Finance*. Kluwer Academic Publishers, 1996.

[7] V. Vapnik. Principles of risk minimization for learning theory. In D.S. Lippman, J.E. Moody, and D.S. Touretzky, editors, *Advances in Neural Information Processing Systems 4*, pp. 831-838. Morgan Kaufmann, 1992.

[8] V. Vapnik. *Statistical Learning Theory*. John-Wiley & Sons, Inc. 1998.

[9] D. Allen. Mean square error of prediction as a criterion for selecting variables. Technometrics, 1971, 13, 469-475.

[10] M. Stone Cross-validatory choice and assessment of statistical predictions. *Journal of Royal Statistical Soc*., 1974, B, 36, pp.111-133.

[11] D.A.Conway, R.R. Marc. Stable factors in security returns: Identification through cross validation. *Journal of Business & Economic Statistics*, 6, January 1988, pp. 1-15.

[12] P. Lachenbruch. An almost unbiased method of obtaining confidence intervals for the probability of misclassification in discriminant analysis. *Biometrics*, 1967, Vol.23, No. 4, pp.639-645.

[13] A. Kostin, O. Krasotkina, M. Markov, V. Mottl, I. Muchnik. Dynamic programming algorithms for analysis of nonstationary signals. *Computational Mathematics and Mathematical Physics*, 2004, Vol. 44, No. 1, pp. 62-77.

[14] M. Markov, O. Krasotkina, V. Mottl, I. Muchnik. Time-varying regression model with unknown time-volatility for nonstationary signal analysis. *The 8th IASTED International Conference on Signal and Image Proces*sing. Honolulu, Hawaii, August 14-16, 2006.

[15] Ph. Jorion. Risk management lessons from long-term capital management. *European Financial Management*, September, 2000.

[16] *Hedge Funds, Leverage, and the Lessons of Long-Term Capital Management*. Report of the President's Working Group on Financial Markets, April 1999, Washington, DC.

[17] *The Wall Street Journal*, 1998-1999.