In proceedings of the 2002 IEEE International Joint Conference on Neural Networks,

May, 2002, Honolulu, Hawaii, USA, 2002

The Prediction Performance of Independent Factor Models

Lai-Wan CHAN
Computer Science and Engineering Department
The Chinese University of Hong Kong, Shatin, HONG KONG
Email : lwchan@cse.cuhk.edu.hk
http://www.cse.cuhk.edu.hk/~lwchan

Abstract - Recently, Independent Component Analysis (ICA) has been proposed to construct factor models in finance. According to the basic principle, the factors extracted using ICA are expected to be independent to each other. This factor model is hence named as independent factor model, in contrast to the traditional factor models which assumes uncorrelated factors. In this paper, we analyze and compare the performance of the independent factor model and tradition factor model based on the prediction ability of the factors. Two examples are given to show that the independent factor model would reduce loss if we have good predictability on one of the factors. On the contrary, uncorrelated factor model may not benefit from an accurate factor prediction.

I. Introduction

A number of well-known financial theories, such as the CAPM and APM, have assumed that the return generation process is governed by the Factor Model [11]. This model describes the return of a security as a weighted linear combination on a number of factors. Equation 1 depicts the Factor Model

$$r_i = \alpha_i + \sum_{m=1}^{k} \beta_{im} F_m + u_i \tag{1}$$

where r_i is the return of security i. k is the number of factors and is a positive integer. $F_1, F_2, ..., F_k$ are the factors affecting the returns of the ith security and β_{i1}, β_{i2}, ..., β_{ik} are the corresponding sensitivities. α_i is a constant named as the "zero" factor. u_i is a zero mean random variable. It is generally assumed that the covariance between u_i and factors F_i are zero. The factors, F_i, are also uncorrelated to each other. u_i and u_j for security i and j are independent if $i \neq j$. In general, the multi-factor model with k factors is called k-factor models.

Traditionally, there are a number of ways to obtain the factor models [9], [4], [8]. We can link factors to some macro-economic measurements, such as unexpected changes in the rate of inflation, interest rate, rate of return on a treasury bill etc. The sensitivities, $\beta's$, are evaluated accordingly. However, the number of factors and which factors should be included are hard to be determined. The other approach is the statistical approach. Principle component analysis (PCA) is the most successful method [7], [10], [12]. It is used to find the factors and their sensitivities[1], [5]. PCA is a suitable tool to construct the factor model because the factors extracted are uncorrelated to each other according to the basic principle of PCA.

However, it has recently been pointed out that uncorrelation is not an appropriate assumption for factor model [6]. Alternatively, it has also been proposed that the factors should be independent [2]. From the viewpoint of portfolio construction, we can construction portfolio to reduce the risk due to certain factors. With uncorrelated factors, it may not be possible to eliminate the factor risk completely. However, with independent factors, it is possible to construct a portfolio which is free from the influence of some of the factors [3].

In addition, there is another advantage of independent factor model over the traditional factor model. In this paper, we show that, by improving the accuracy in the prediction of the factors, the possibility of having a loss would be reduced when the return is formed by the independent factor model. However, the loss may not be reduced in the case of uncorrelated factor model. We will give examples in the following sections to illustrate that uncorrelated factor model may not be able to gain any benefit in terms of return prediction when the accuracy in factor prediction increases.

II. The Prediction Error in the Factor Model

For simplicity, we consider a stock governed by a 2-factor model as shown below.

$$r = \alpha + \beta_1 F_1 + \beta_2 F_2 + u \tag{2}$$

Let $\hat{F}_1$ and $\hat{F}_2$ be the predicted values of factors F_1 and F_2 respectively. We put ϵ_1 and ϵ_2 as the prediction errors of these two factor terms, *i.e.* $\epsilon_1 = \beta_1 F_1 - \beta_1 \hat{F}_1$, and $\epsilon_2 = \beta_2 F_2 - \beta_2 \hat{F}_2$. In this case, the error of the return due to the error in prediction of the factors is $\epsilon_r = \epsilon_1 + \epsilon_2$.

If $p_1(\epsilon_1)$ and $p_2(\epsilon_2)$ denote the probability density function of ϵ_1 and ϵ_2, the cumulative distribution function of ϵ_r would be

$$\begin{aligned} P_r(\epsilon_r) &= \int_{-\infty}^{\epsilon_r} p_r(z)dz \\ &= \int_{\epsilon_2=-\infty}^{\epsilon_r - x} \int_{\epsilon_1=-\infty}^{x} p_r(\epsilon_1 \wedge \epsilon_2) d\epsilon_1 d\epsilon_2 \\ &= \int_{\epsilon_2=-\infty}^{\epsilon_r - x} \int_{\epsilon_1=-\infty}^{x} p_1(\epsilon_1) p_2(\epsilon_2 | \epsilon_1) d\epsilon_1 d\epsilon_2 \end{aligned} \quad (3)$$

Equation 3 denotes the probability of the downside error in the return due to the predicted errors in the factors, *i.e.*, probability of the return which falls below ϵ_r of the predicted return.

In financial trading, decisions would often be made according to the predicted values of the stocks. One major concern is on the analysis of the risk of our decision, especially the downside risk. Therefore, equation 3 provides an expression for us to analyze the risk of a predicted value of the stock, and it relates the error of the return to the errors on the prediction of the factors.

If ϵ_1 and ϵ_2 are independent, equation 3 turns into

$$P_r(\epsilon_r) = \int_{\epsilon_2=-\infty}^{\epsilon_r - x} \int_{\epsilon_1=-\infty}^{x} p_1(\epsilon_1) p_2(\epsilon_2) d\epsilon_1 d\epsilon_2 \quad (4)$$

We can see that the density function of $p_r(\epsilon_r)$ is the convolution of $p_1(\epsilon_1)$ and $p_2(\epsilon_2)$. When we have more than two independent factors in our model, the same technique can be applied to equation 4 to further generalize the equation to include more factors. In addition, if we want to incorporate the effect due to u, which is also independent to the factors, we can convolute the error density with the density of u.

III. The First Example

We use an example to illustrate the difference between having independent factors and non-independent factors on the prediction error of the return. In case A, ϵ_1 and ϵ_2 are independent to each other. It is reasonable to assume that the factor errors following the Gaussian distribution. We generate 10,000 random samples, following Gaussian distribution with zero mean and unit variance, for both

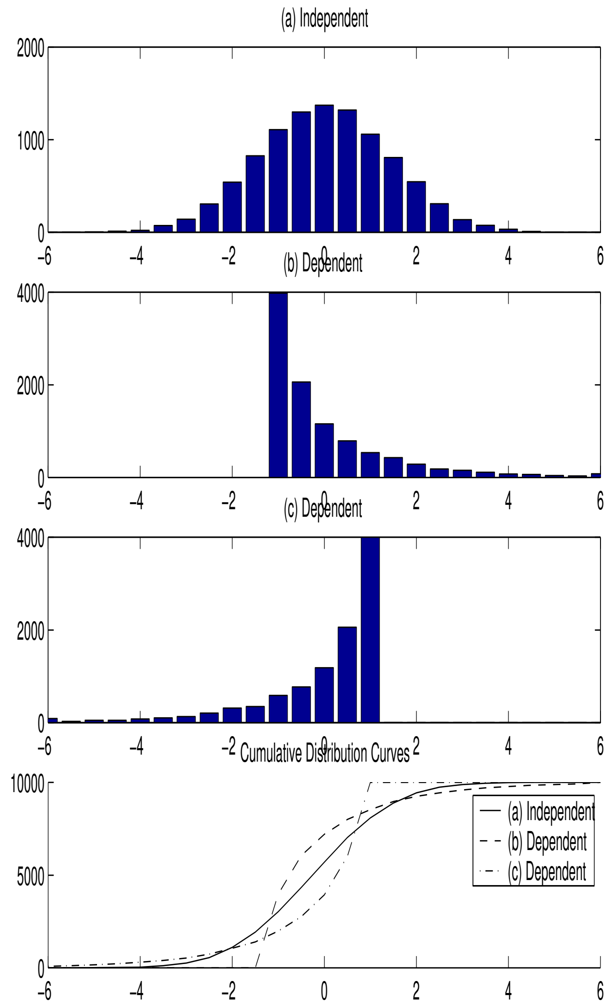

Fig. 1. The density of ϵ_r in three cases; (a) ϵ_1 and ϵ_2 are independent, (b) $\epsilon_2 = \frac{\epsilon_1^2 - 1}{1.4}$, (c) $\epsilon_2 = \frac{1-\epsilon_1^2}{1.4}$, (d) the cumulative distribution curves of the three cases.

variables ϵ_1 and ϵ_2. Figure 1(a) shows the distribution of their sum, which is also following Gaussian distribution.

In case B, the two factors are dependent. Suppose $\beta_2 F_2 = \frac{1}{a}(\beta_1 F_1)^2 + b$ where a and b are two constants. In this case, although the factors are dependent, they are uncorrelated to each other. With such a relationship, we obtain $\epsilon_2 = -\frac{2}{a}\epsilon_1 \beta_1 F_1 + \frac{1}{a}\epsilon_1^2$. F_1 is involved in the first term of this expression. With the fact that the factors are zero-mean and the assumption that the predicted error, ϵ_1, is independent to the values of the factor, the expected values of the first term over F_1 would be vanished. This leaves $\epsilon_2 \propto \epsilon_1^2$. It is interesting to note that a zero-mean ϵ_1 does not imply a zero-mean ϵ_2. This leads to a bias in the prediction error of F_2 and hence the return.

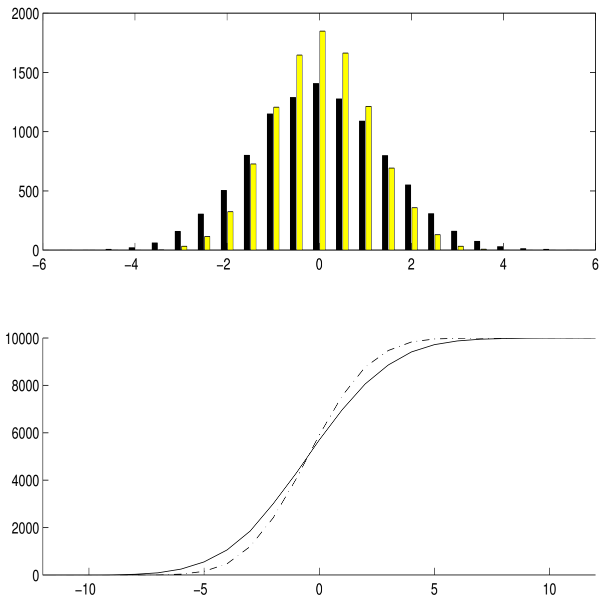

Fig. 2. Case A. The density and cumulative distribution of ϵ_r when the variance of ϵ_1 and ϵ_2 has been changed. The dark bars in the top figure and the solid lines in the bottom figure indicate the case when variance = 1. The light bars and dotted lines indicate the case when variance = 0.75.

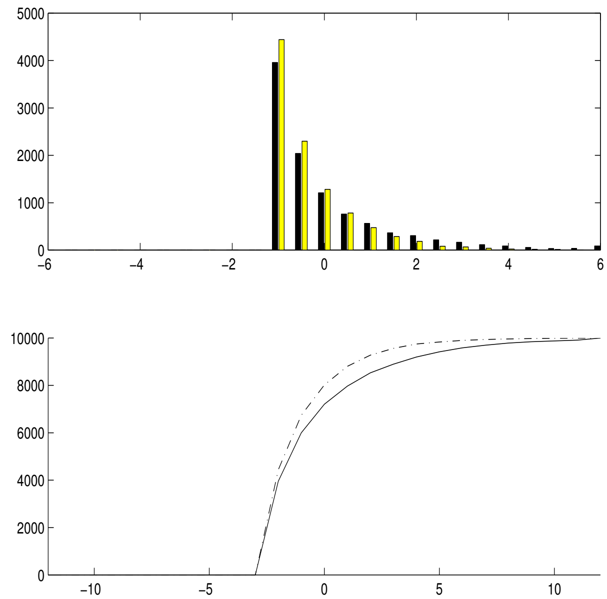

Fig. 3. Case B. The density and cumulative distribution of ϵ_r when the variance of ϵ_1 has been changed. The dark bars in the top figure and the solid lines in the bottom figure indicate the case when variance = 1. The light bars and dotted lines indicate the case when variance = 0.75.

To make the expected error in the return comparable to the case in the independent case, we remove the bias by adding a constant term to ϵ_2. Let $\epsilon_2 = \frac{\epsilon_1^2-1}{1.4}$. The constants in this expression are assigned such that ϵ_2 has the same mean and variance as in the independent case (*i.e.* mean = 0 and variance = 1). Now in both case A and case B, ϵ_1 and ϵ_2 are uncorrelated to each other. Figure 1(b) shows the distribution of ϵ_r. It can be seen that the distribution in this figure is significantly different from the case with independent variables.

In the third case, C, ϵ_2 is defined as $\epsilon_2 = \frac{1-\epsilon_1^2}{1.4}$. Similar to case B, the factors F_1 and F_2 are uncorrelated but not independent, but a is negative in case C. The distribution of ϵ_r in case C appear in a different form (see Figure 1(c)). Figure 1(d) compares the cumulative distribution function of three cases. In the dependent cases (B and C), it is possible to make $P(\epsilon_r)$ either larger or smaller than that in the independent case. This relies on how the factors are dependent on each other. In all cases, the two factors are uncorrelated to each other, but only case A has independent factors. When the variables are added together, the same mean, and as well as the same variance, of the sums are obtained in all cases (*i.e.* mean = 0, variance = 2). The only difference is on the probability distribution of the sums. In the next section, we will show how these distributions affect the errors of the predicted returns.

IV. The Prediction Error in the First Example

Suppose we have a predictor which gives a better prediction on the factors. For a good predictor, we refer to one which gives a smaller error variance in the prediction of the factors (*i.e.* variance of ϵ_1 and ϵ_2 are smaller). For example, when we reduce the variance of the errors in each component to 0.75, Figure 2 shows the density and cumulative distribution function of ϵ_r in case A, compared to that when the variance is equal to 1. It shows that $P_r(\epsilon_r)$ has been reduced when $\epsilon_r < 0$. However, in the dependent cases, case B (Figure 3) shows an increase in $P_r(\epsilon_r)$ for $\epsilon_r < 0$ while case C (Figure 4) shows a decrease in $P_r(\epsilon_r)$ for $\epsilon_r < 0$. In conclusion, the independent factors would reduce the probability of the downside error in the return when we have a good predictor. In the cases of the dependent factors, the downside error could be increased or decreased, depending on the relationship between the factors.

V. The Second Example

The example in the previous section compares the independent factors and the dependent factors. In this section, we demonstrate an example which requires the

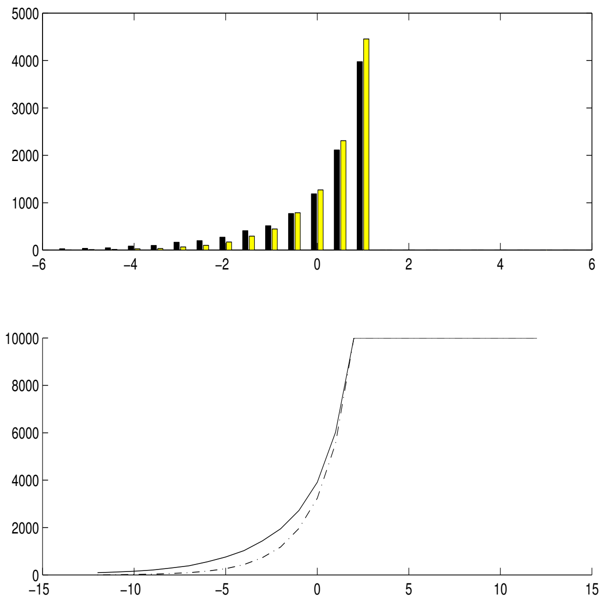

Fig. 4. Case C. The density and cumulative distribution of ϵ_r when the variance of ϵ_1 has been changed. The dark bars in the top figure and the solid lines in the bottom figure indicate the case when variance = 1. The light bars and dotted lines indicate the case when variance = 0.75.

decomposition of the factors, either using independent factor model or traditional factor models. We would also show the effect of having a good predictor in both models.

Suppose the distribution of two stock prices are given as in Figure 5, *i.e.*uniformly distributed within the rhombus. The x-axis and y-axis of the figure denote the prices of the stocks. When we construct the factor model from the data, it is possible to obtain an infinite number of factor models. In traditional factor models, we require that the factors are uncorrelated. With this uncorrelation restriction, it is still possible to obtain more than one solution. For example, Figure 6 shows the axes of two uncorrelated factors, x_1 and x_2, denoted by the two straight lines on the graph. These factors are obtained using Principle Component Analysis (PCA). The axes are extracted in a way that the error variance along the principle axis is minimized. It has to be noted that the axes are orthogonal to each other. Although the factors are uncorrelated, they are not independent. Figure 7 shows the distribution of another set of uncorrelated factors, y_1 and y_2. obtained by Independent Component Analysis (ICA). In this case, the factors are independent and uncorrelated. In order to make the factors independent, the axes are no longer orthogonal.

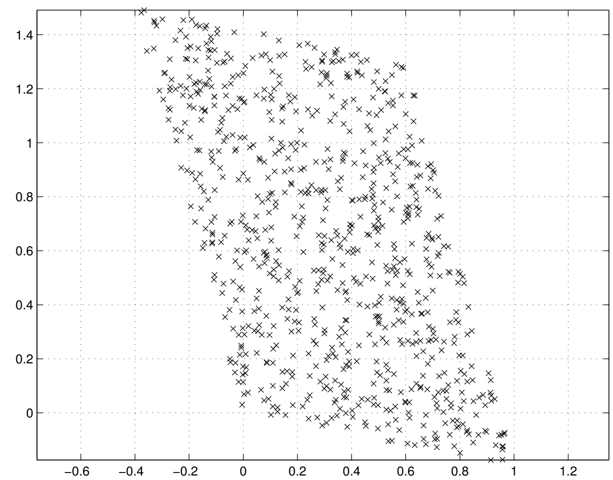

Fig. 5. The distribution of the returns of two stocks

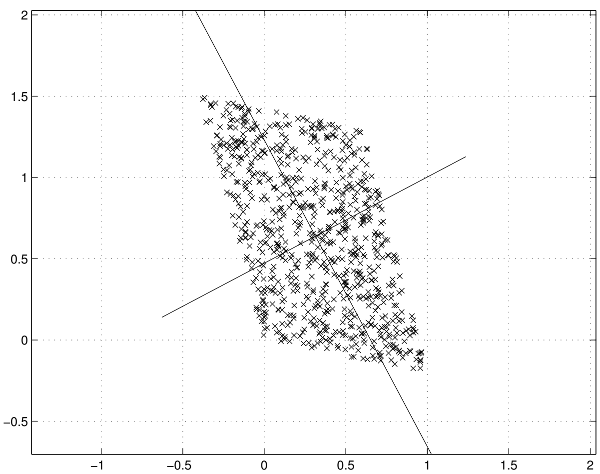

Fig. 6. The lines indicate the axes of the uncorrelated factors extracted using PCA.

VI. The Prediction Error in the Second Example

Let us assume that there is a good predictor for the factors in the factor model. If there is such a predictor, we could use the predicted factors to estimate the stock returns and hopefully we could reduce the risk of having unexpected poor returns. Similar to the previous example, we define that a good predictor is one which reduces the variance of the errors. In the subsequent figures, we use dots to denote the outcomes of the uniformly distributed stock returns, and the circles to denote the difference between the distribution in Figure 5 and the outcomes from our good predictor. In other words, the circles indicate those unexpected poor returns that we

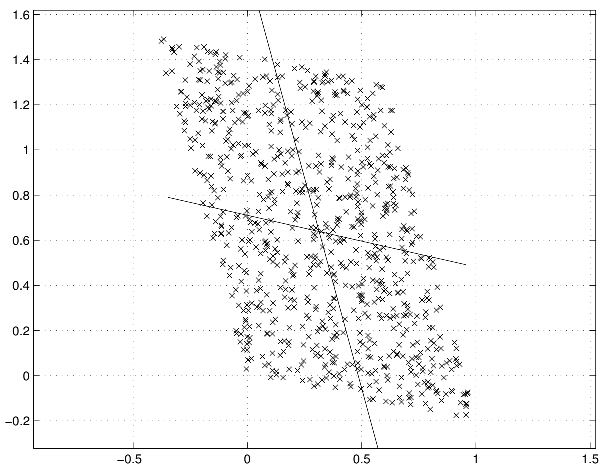

Fig. 7. The lines indicate the axes of the independent factors extracted using ICA.

want to avoid.

In the case of independent factor model, if a good predictor helps us to reduce the error of one of the factors, (indicated by the circles in Figure 8), this knowledge is very useful for us work out the distribution of the stock returns. Projecting the data into the x-axis gives us the distribution of the stock values of the first stock. As the circles are concentrated on the left hand side, it is easy to figure out that we have substantially avoid the chances of have a a low return on the first stock. Similarly, with a good predictor on the other factor, (Figure 9), we could avoid certain low returns on the second stock, with values projected on the y-axis.

If the factor model is constructed by PCA, we may not gain the same benefit. Suppose we have knowledge to avoid the low return in one of the factors (indicated by the circles in Figure 10). Similar to the independent factor model, we can make use this knowledge to reduce the downside risk of the first stock. On the other hand, with a good predictor of the second factor to avoid certain negative errors in the second factor (indicated by the circles in Figure 11), it is impossible for us to avoid the occurrence of the lowest returns of neither stocks. This example shows that a good predictor in one or some of the factors extracted by the factor models using ICA may not help us to have a better prediction on the stock returns. This is different from the independent factor model, which the distribution of the return is the convolution of the distributions of the individual factors. Hence, we would certainly be able to benefit from a good predictor on the independent factors.

VII. Conclusions and Discussions

In this paper, by giving two examples, we have demonstrated the difference between factor models using independent factors and uncorrelated factors. Independent factors imply uncorrelated factors but not vice versa. One advantage of the independent factor models is that the resultant distribution of the return is the convolution of the individual factors. Analysis of the return is relatively easy, comparing to the dependent cases. In terms of prediction performance, independent factor models would gain advantage in the estimation of the return and in the avoidance of the loss when having a more accurate prediction on the factors. This, however, may not be the case in uncorrelated factor models.

VIII. Acknowledgement

The work described in this paper was partially supported by a grant from the Research Grants Council of the Hong Kong Special Administration Region, China.

References

[1] S. Brown. The number of factors in security returns. *The Journal of Finance*, 44, no. 5:1247–1262, December 1989.

[2] S.M. Cha and L.W. Chan. Applying independent component analysis to factor model. In L.W. Chan K.S. Leung and H. Meng, editors, *Intelligent Data Engineering and Automated Learning - IDEAL 2000, Data Mining, Financial Engineering and Intelligent Agents*, pages 538–544. Springer, 2000.

[3] L.W. Chan and S.M. Cha. Selection of independent factor model in finance. In *The Proceedings of the Third International Conference on Independent Component Analysis and Signal Saparation, ICA2001, San Diego, U.S.A.*, 2001.

[4] N.F. Chen, R. Roll, and S. Ross. Economic forces and the stock market. *Journal of Business*, 59, no. 3:383–403, July 1986.

[5] G. Connor and R. Korajczyk. Performance measurement with the arbitrage pricing theory a new framework for analysis. *Journal of financial economics*, 15:373–394, 1986.

[6] P. Embrechts, A.J. McNeil, and D. Straumann. Correlation and dependence in risk management: Properties and pitfalls. In M. Dempster, editor, *to appear in RISK Management: Value at Risk and Beyond*. Cambridge University Press, 2001.

[7] G. Feeney and D. Hester. Stock market indices: A principal component analysis. *Cowles Foundation*, Monograph 19(39):110–138, 1967.

[8] A. Gordon, W. Sharp, and B. Jeffery. *Fundamentals of investments*. Prentice Hall, 1993.

[9] B. Manly. *Multivariate statistical methods: A primer*. Chapman and Hall, 1994.

[10] H. Schneeweiss and H. Mathes. Factor analysis and principal components. *Journal of multivariate analysis*, 55:105–124, 1995.

[11] W. F. Sharpe. *Investments*. Prentice-Hall, 1981.

[12] J. Utans, W.T. Holt, and A.N. Refenes. Principal components analysis for modeling multi-currency portfolios. In *Proceedings of the Fourth International Conference on Neurals Networks in the Capital Markets, NNCM-96*, 1997.

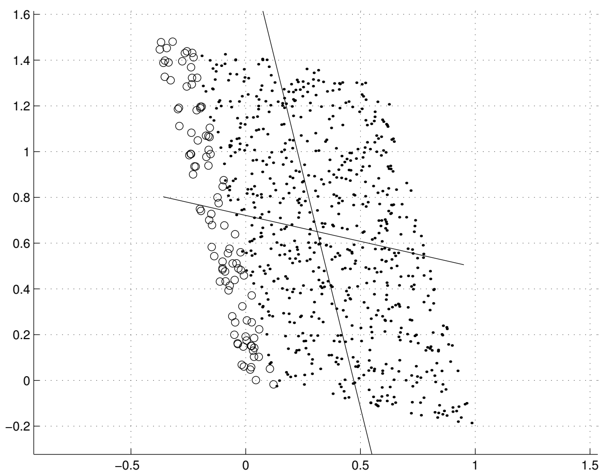

Fig. 8. Independent factors extracted using ICA. The circles indicate the distribution of low returns on factor 1.

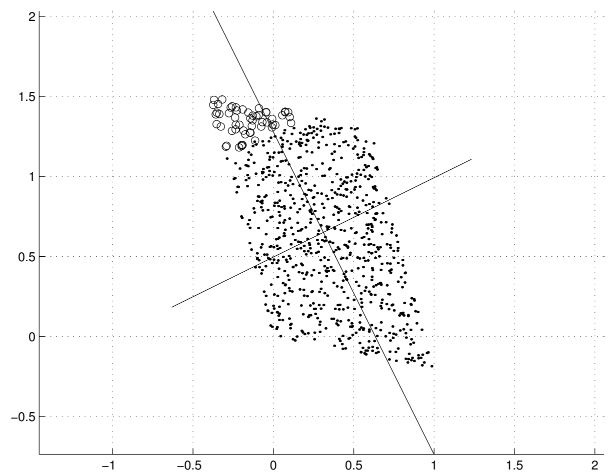

Fig. 10. Uncorrelated factors extracted using PCA. The circles indicate the distribution of low returns on factor 1.

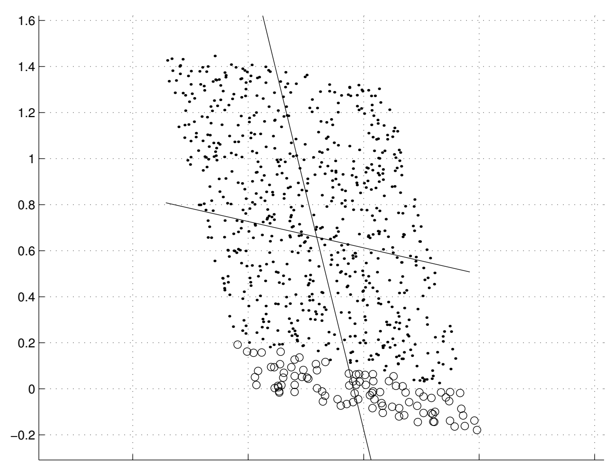

Fig. 9. Independent factors extracted using ICA. The circles indicate the distribution of low returns on factor 2.

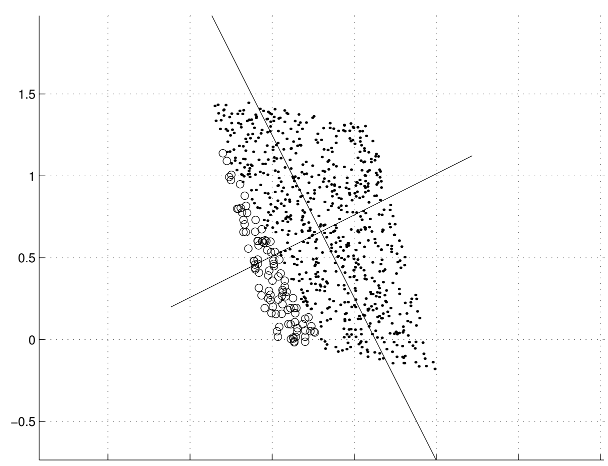

Fig. 11. Uncorrelated factors extracted using PCA. The circles indicate the distribution of low returns on factor 2.