Statistical Model

Fixed Income Derivatives

Yuri Balasanov

MSFM, University of Chicago

Disclaimer

The content of these lectures, any comments, opinions and materials are put together by the author especially for the course identified in the title, they are sole responsibilities of the author, but not of the author's employers or clients.

The author cannot be held responsible for any material damage as a result of use of the materials presented in this document or in this course.

For any inquiries contact the author, Yuri Balasanov, at
ybalasan@uchicago.edu or yuri.balasanov@research-soft.com

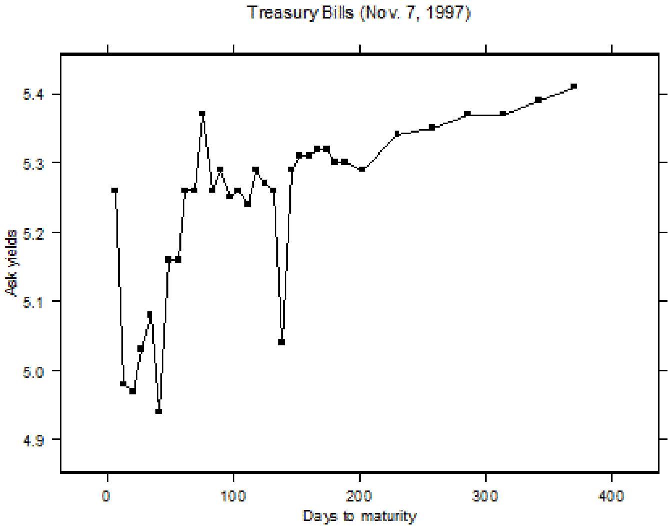
Outline of This Lecture

- Review of PCA
- Problem of reduction of dimensionality
- Data for description
- PCA algorithm for interest rates
- Interpretation of the results
- Turning data description into a model of interest rates
- Application to risk management
- Homework assignment

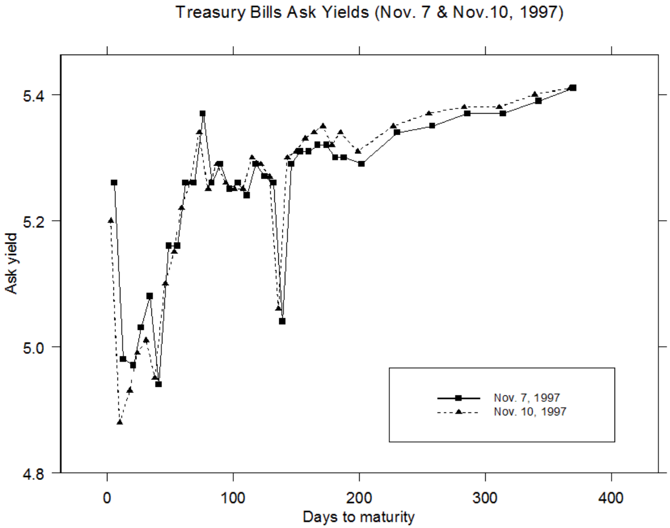
Generations of Interest Rates Models

1. 1980s; First generation: One-factor models;
   - Had analytical formula for discount bond and option on bond
   - Not all of them were arbitrage-free (CIR)
   - They could not generate term curves and rates correlation structures consistent with observations
2. 1990s; Second generation: Multifactor HJM
   - General framework for constructing interest rate models; Allows for multiple stochastic drivers, which improves the correlation structure for rates with different maturities
   - Includes majority of the one-factor models as particular cases
   - Analytical formulas for bonds and bond options
   - Difficult to calibrate
3. 2000-current; Third generation: market models;
   - Solved the problem of calibration
   - Correlation structure for rates of different maturities needs to be estimated separately
   - In fact belong to the HJM framework

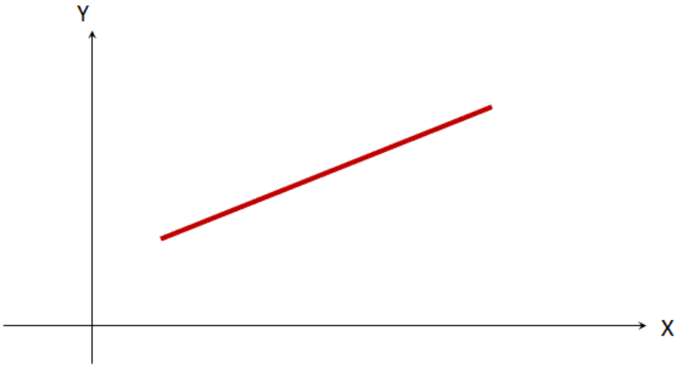
Treasury Yields to Maturity: Day 1



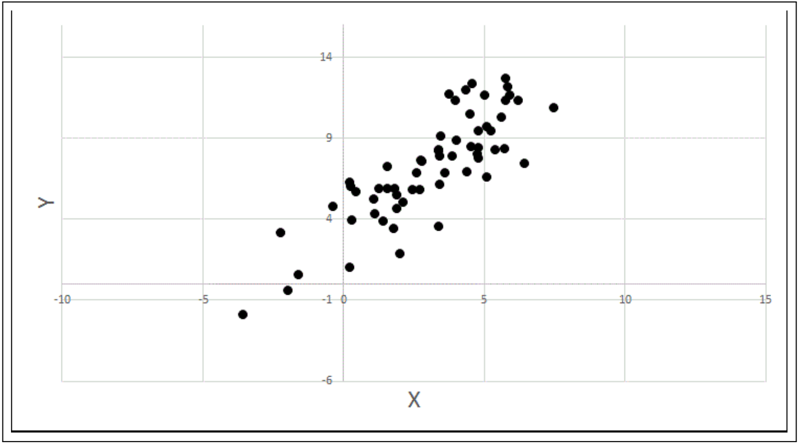
Treasury Yields to Maturity: Day 2



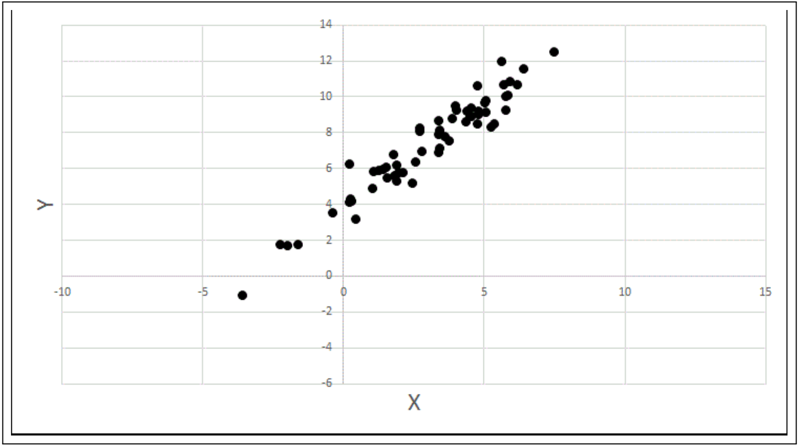
How Many Dimensions Data Have?



How Many Dimensions Data Have?



How Many Dimensions Data Have?



How Can We Take Advantage of Low Dimension?

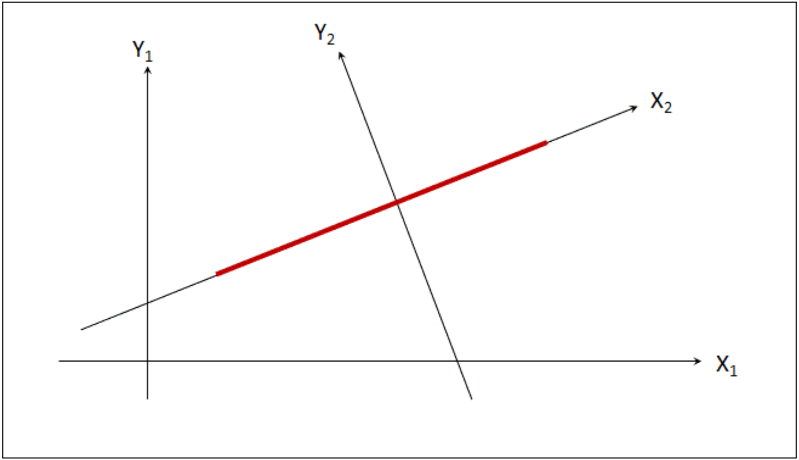


Figure: Even if we know that a line has dimension of one it is still a collection of points on the plane... Until we change the coordinates

Finding the Right Coordinates

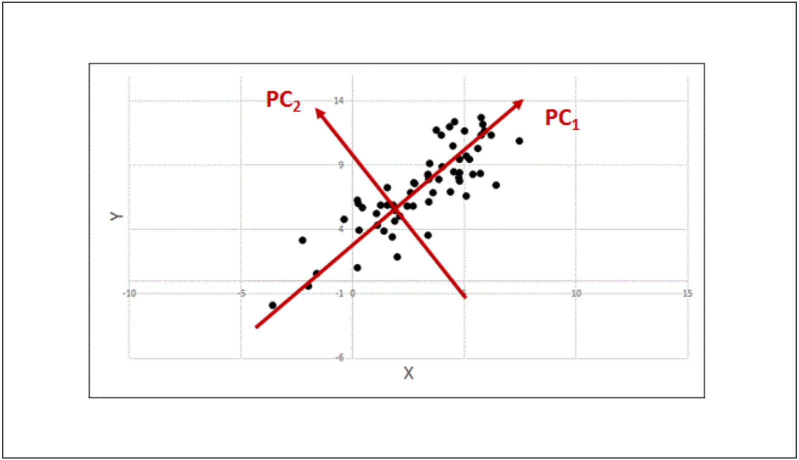


Figure: Finding a set of new orthonormal vectors forming the new coordinates of the reduced dimensionality space is the objective of PCA.

Matrix Decomposition

Let $\mathbf{u}_1, \ldots, \mathbf{u}_n$ be linearly independent (orthogonal) eigenvectors corresponding to the eigenvalues $\lambda_1, \ldots, \lambda_n$ of a $n \times n$ matrix A. Each eigenvector is a column $\mathbf{u}_i = \langle u_{1i}, \ldots, u_{ni} \rangle^T$.

Fact

If A is a symmetric matrix with real values then it has n real eigenvalues (not necessarily all different) corresponding to n eigenvectors $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n$ forming an orthonormal basis of $\mathbb{R}^n$.

Fact

If A is a real symmetric matrix with eigenvalues $\lambda_1, \ldots, \lambda_n$ and eigenvectors $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n$ then

$$A = U \Lambda U^T = [\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n] \begin{bmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{bmatrix} [\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n]^T,$$

Spectral Decomposition of a Matrix

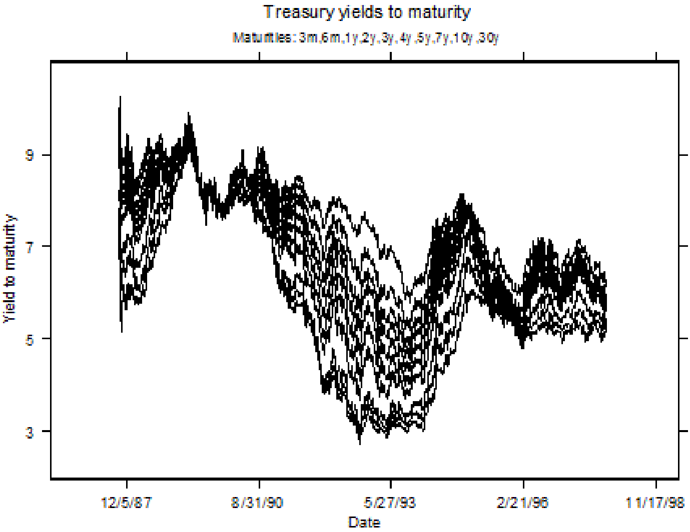
Fact

If A is a real symmetric matrix with eigenvalues $\lambda_1, \ldots, \lambda_n$ and eigenvectors $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n$ then

$$A = \sum_{i=1}^{n} \lambda_i \mathbf{u}_i \mathbf{u}_i^T.$$

Decomposition does not depend on the order of eigenvalues and eigenvectors, as long as they are in the same order.
We will assume that eigenvalues are in decreasing order.
Spectral decomposition of a symmetric real-valued matrix A shows that it is a sum of n components, each defined by the corresponding eigenvector (direction). And the weight of each component in the sum is the corresponding eigenvalue.
Components with small contributions (small eigenvalues) can be ignored to reduce the dimensionality.

Data for PCA I



Data for PCA II

Let **Y** be the matrix of yields with maturities from 3 months to 30 years (no Output column).

We would like to find approximation $\hat{\mathbf{Y}}$ to **Y** of the form

$$\begin{aligned} \hat{\mathbf{Y}}_0 &= \mathbf{F}\mathbf{L}^T, \\ \hat{\mathbf{Y}}_0(t,\tau) &= \hat{\mathbf{Y}}(t,\tau) - \mathbf{L}_0(\tau) = \sum_{k=1}^{m} l_i(\tau) f_k(t), \\ \mathbf{L}_0(\tau) &= \mathbb{E}\left[\mathbf{Y}(\cdot,\tau)\right], \end{aligned}$$

such that $\hat{\mathbf{Y}}$ has as close covariance matrix to that of **Y** as possible.

Columns of **F** are called **factors** or **factor scores**. Rows of $\mathbf{L}^T$ are called **factor loadings**.

New factor space is m-dimensional. Choice of m is based on quality of fit.

Covariance matrix of **Y** is real-valued symmetric with positive eigenvalues.

Example

Calculate covariance matrix of the yields data.

Outline of the Algorithm

1. Create centered matrix

$$\begin{aligned}\mathbf{Y}_0 &= \left[\mathbf{y}^0_{\cdot 1}, \mathbf{y}^0_{\cdot 2}, \ldots, \mathbf{y}^0_{\cdot n}\right] \\ &= \left[\mathbf{y}_{\cdot 1} - \mathbf{1}\bar{y}_{\cdot 1}, \mathbf{y}_{\cdot 2} - \mathbf{1}\bar{y}_{\cdot 2}, \ldots, \mathbf{y}_{\cdot n} - \mathbf{1}\bar{y}_{\cdot n}\right]\end{aligned}$$

2. Calculate covariance matrix

$$\Theta = cov\left(\mathbf{Y}_0\right) = \left\{\mathbb{E}\left[\mathbf{y}^0_{\cdot i}\mathbf{y}^0_{\cdot j}\right]\right\}$$

3. Perform eigenvalue decomposition of Θ. Define

$$\mathbf{L} = \left[\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m\right],$$

where $\mathbf{u}_i$ is the eigenvector corresponding to the i-s largest eigenvalue of Θ.

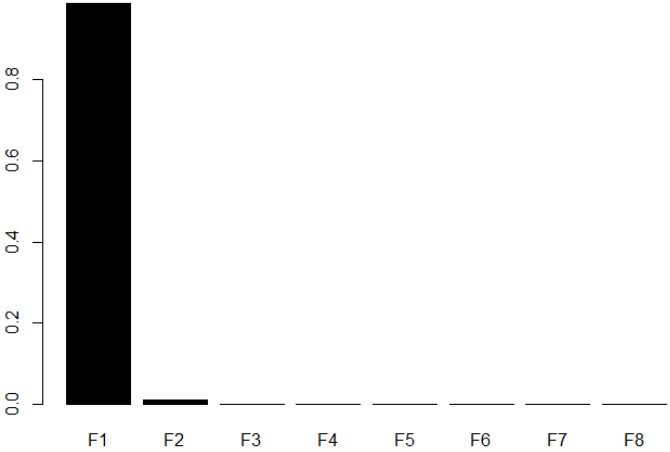
4. Define

$$\mathbf{F} = \mathbf{Y}_0\mathbf{L},$$

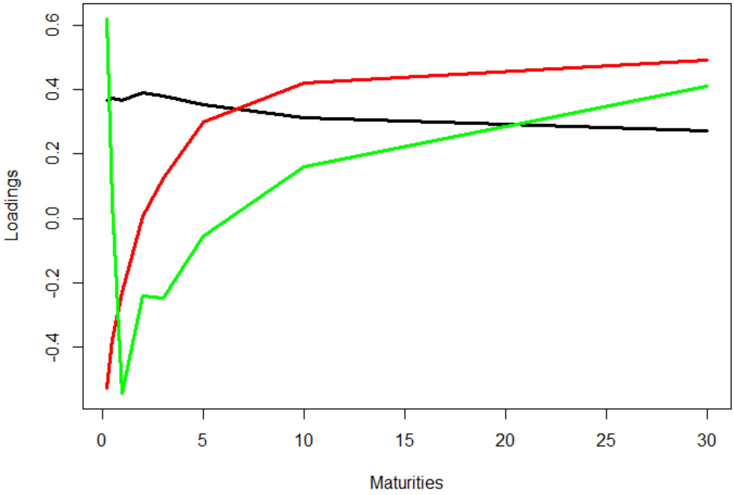
right-multiplying by $\mathbf{L}$ the definition of the model

$$\mathbf{Y}_0 = \mathbf{F}\mathbf{L}^T.$$

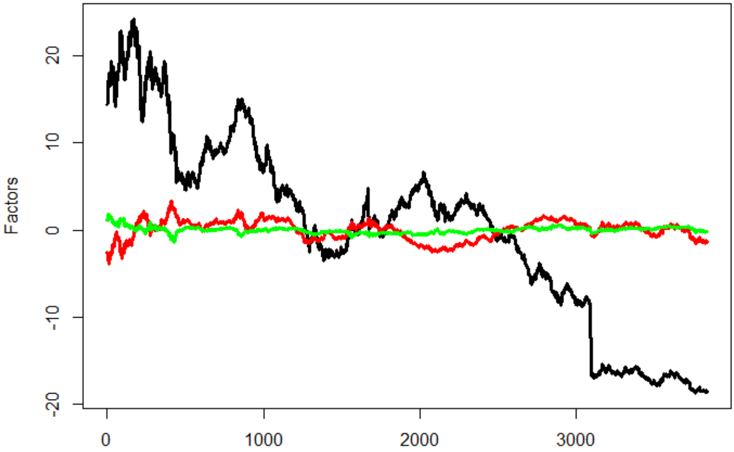
Why PCA Approximation Works?



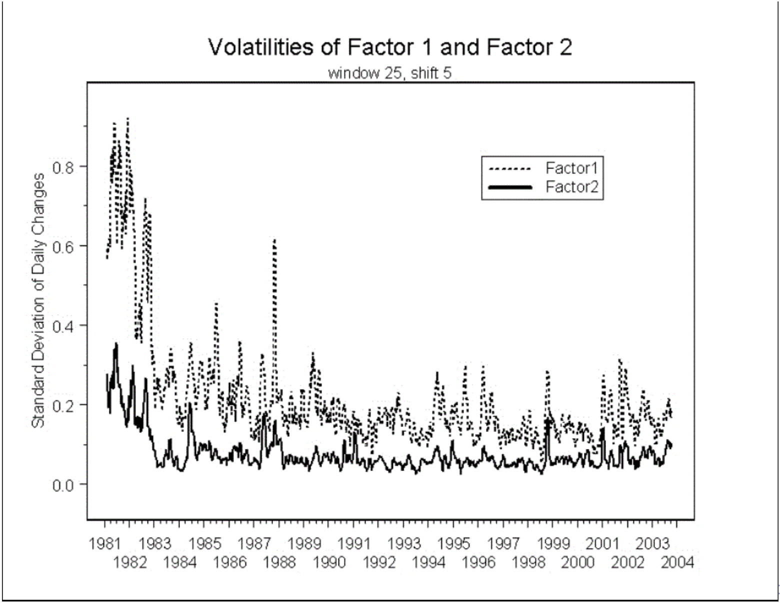
Factor Loadings



Factors



Factor Volatilities



Interpretation of Factors and Loadings

Estimated model for $\tau = 0.25, 0.5, 2, 3, 5, 10, 30$ is:

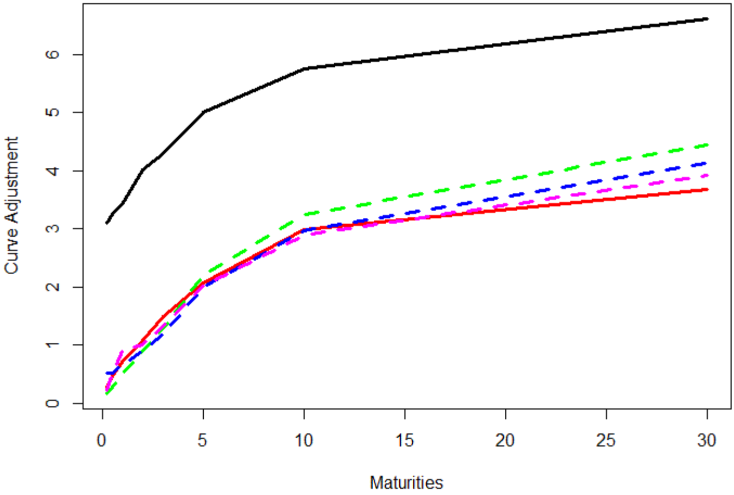
$$\hat{y}(t, \tau) = l_0(\tau) + l_1(\tau) f_1(t) + l_2(\tau) f_2(t) + l_3(\tau) f_3(t).$$

Look at the change in the curve $\hat{y}(t, \tau)$ from day to day:

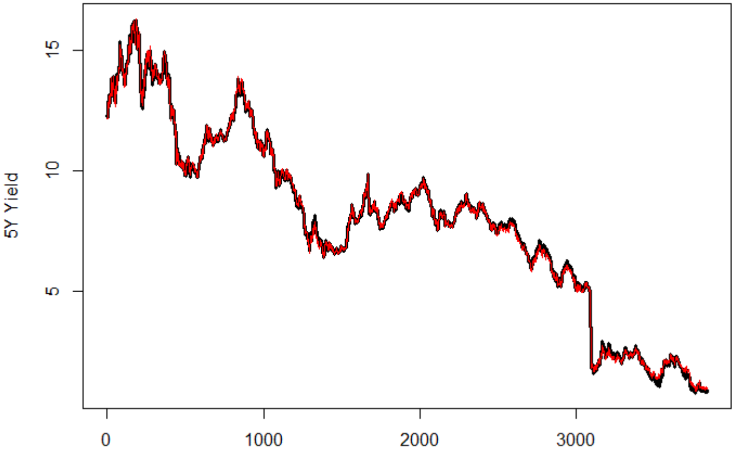
$$\begin{aligned} \hat{y}(t+1, \tau) - \hat{y}(t, \tau) &= l_1(\tau)(f_1(t+1) - f_1(t)) \\ &+ l_2(\tau)(f_2(t+1) - f_2(t)) \\ &+ l_3(\tau)(f_3(t+1) - f_3(t)) \end{aligned}$$

Remember that loadings do not change with time.
Each of the factors makes an adjustment corresponding to the shape of its loading.
Look at the shapes of the loadings and tell what mode of curve move corresponds to each factor.

Adjustment of the Curve by Each of the 3 Factors



Quality of Fit



Most Remarkable Properties of Factor Decomposition

- Convenient factorization. Our parameters are divided into 2 groups. One, factor loadings, depends only on maturity, but not on time; the other, factors, depends only on time, but not on maturity
- Accuracy. The maximum deviation of our approximation from the original data is within 50 basis points over the period of more than 20 years of daily observations
- Low dimensionality. Recall that factor loadings do not depend on time and remain constant through the whole period. Then on each day the term curve is define by just 3 numbers, the factors
- Interpretation of the factors as the 3 modes of move
- Possibility of turning this statistical description of the data into an interest rate model of HJM type

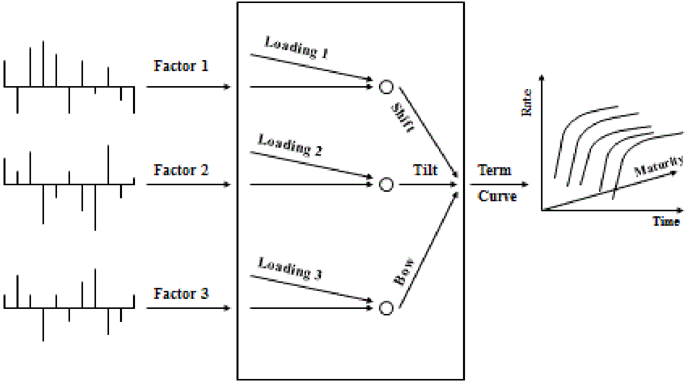
Inhibit Wrong Neural Binding

- Common statement in many textbooks:

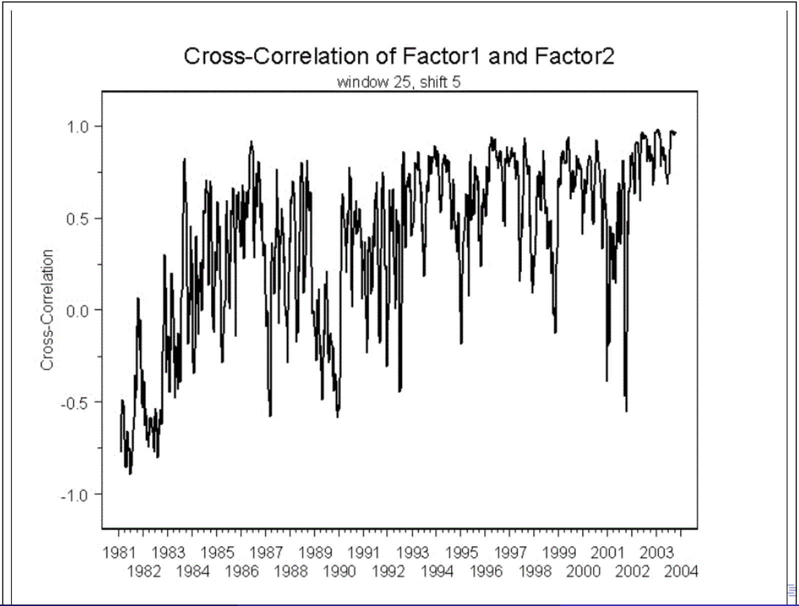
it is a mathematical procedure that uses an orthogonal transformation to convert a set of observations of possibly correlated variables into a set of values of uncorrelated variables called principal components

- This stereotype property of the PCA method is misleading.
- Vectors of factorsfor the whole period of observation are linearly independent by construction
- This is different from zero correlation between two stochastic processes
- Correlation structure changes in time and can be interpreted

Interpretation of PCA Model



Correlation Between Principal Components



The model expression is:

$$R\left(t,\tau\right)=L_0\left(\tau\right)+\sum_{i=1}^{3} l_i\left(\tau\right)\times f_i\left(t\right)$$

where , $L_0\left(\tau\right), l_1\left(\tau\right), l_2\left(\tau\right), l_3\left(\tau\right)$ are factor loadings and $f_1\left(t\right), f_2\left(t\right), f_3\left(t\right)$ are factors.
The forward rate by definition is

$$\begin{aligned} F\left(t,\tau_1,\tau_2\right) &= \frac{R\left(t,\tau_2\right)\tau_2-R\left(t,\tau_1\right)\tau_1}{\tau_2-\tau_1} \\ &= \frac{\left(B_0\left(\tau_2\right)-B_0\left(\tau_1\right)\right)}{\tau_2-\tau_1}+\sum_{i=1}^{3}\frac{\left(B_i\left(\tau_2\right)-B_i\left(\tau_1\right)\right)}{\tau_2-\tau_1}\times f_i\left(t\right) \end{aligned}$$

where

$$B_i\left(\tau\right)=l_i\left(\tau\right)\tau$$

The instantaneous forward rate is

$$
\begin{aligned}
F(t,\tau) &= \lim_{\tau_1\to\tau} F(t,\tau,\tau_1) \\
&= \lim_{\tau_1\to\tau}\left(\frac{(B_0(\tau_1)-B_0(\tau))}{\tau_1-\tau}+\sum_{i=1}^{3}\frac{(B_i(\tau_1)-B_i(\tau))}{\tau_1-\tau}\times f_i(t)\right) \\
&= B_0'(\tau)+\sum_{i=1}^{3}B_i'(\tau)\times f_i(t)
\end{aligned}
$$

and instantaneous spot rate is

$$
R(t)=\lim_{\tau\to 0}F(t,\tau)=B_0'(0)+\sum_{i=1}^{3}B_i'(0)\times f_i(t)
$$

Language of Factors and Loadings

Bond Prices

The zero-coupon bond price corresponding to the rate $R(t,\tau)$ is

$$
\begin{aligned}
P(t,\tau) &= \exp\left(-R(t,\tau)\,\tau\right) \\
&= \exp\left(-\left(B_0(\tau) + \sum_{i=1}^{3} B_i(\tau) \times f_i(t)\right)\right)
\end{aligned}
$$

The forward bond price is

$$
\begin{aligned}
P(t,\tau_1,\tau_2) &= \exp\left(-F(t,\tau_1,\tau_2)(\tau_2-\tau_1)\right) \\
&= \exp\Bigg((B_0(\tau_1) - B_0(\tau_2)) \\
&\quad + \sum_{i=1}^{3} (B_i(\tau_1) - B_i(\tau_2)) \times f_i(t)\Bigg)
\end{aligned}
$$

Now note that spot and forward rates and logarithms of spot and forward bond prices are linear functions of factors f_i.

Units

All rates are proportional to one over time

$$R(t,\tau), F(t,\tau_1,\tau_2), F(t,\tau) \propto \frac{1}{t}$$

The random factor is measured by the same unit

$$f_i \propto \frac{1}{t}$$

Loadings have different units

$$\begin{array}{l} L_0(\tau) \propto \frac{1}{t}, \\ L_i(\tau) \propto 1, i \geq 1, \\ B_0(\tau) \propto 1, \\ B_i(\tau) \propto t, i \geq 1. \end{array}$$

Parametric Assumptions About Loadings

Turning statistical model into a mathematical model of interest rates requires replacement of eigenvectors with a parametric loading curves. Commonly used (Heath-Jarrow-Morton, Vasicek, Hull-White models) parametric forms for $l_1(\tau)$ in one-factor case is $l_1(\tau) = \frac{1-\exp(-a\tau)}{a\tau}$. With such function single factor is a linear function of the short rate:

$$
\begin{aligned}
B_1'(\tau) &= \left[\frac{1-\exp(-a\tau)}{a}\right]' = \exp(-a\tau), \\
F(t,\tau) &= B_0'(\tau) + \exp(-a\tau) \times f_1(t), \\
R(t) &= B_0'(0) + f_1(t).
\end{aligned}
$$

We use a more general form suggested in [1]

$$
l_i(\tau) = \sum_{j=1}^{m_i} b_{ij} \frac{1-\exp(-a_{ij}\tau)}{a_{ij}\tau}
$$

O. Cheyette, 1994, Markov Representation of the Heath-Jarrow-Morton Model, Working paper.

Factors as Stochastic Drivers

The industry standard assumptions about stochastic dynamics of factors are the two main types of stochastic differential equations:

1. Simple diffusion

$$df_i(t) = \sigma_i dw_i(t)$$

2. Ornstein-Uhlenbeck process

$$df_i(t) = -af_i(t)\,dt + \sigma_i dw_i(t)\,.$$

In both formulas w_t is a Weiner process and a, σ_i are the mean reversion and volatility parameters.
The discrete analog of this process is

$$f_i(t) = (1-a)\,f_i(t-1) + \sigma_i \varepsilon.$$

which is in fact is a first order autoregression process.
We also assume for simplicity that the three factors are uncorrelated. One can always rotate the vector of Gaussian factors to make them uncorrelated.

Correlations Structure of Forward Rates

Using the definition of instantaneous forward rates find correlation between $\Delta F\left(t, t+\tau_1\right)$ and $\Delta F\left(t, t+\tau_2\right)$.

$$\rho\left(\Delta F\left(t,\tau_1\right), \Delta F\left(t,\tau_2\right)\right) = \frac{\sum_{i=1}^{3} B_i'\left(\tau_1\right) B_i'\left(\tau_2\right) \sigma_i^2}{\sqrt{\left[\sum_{i=1}^{3}\left(B_i'\left(\tau_1\right)\right)^2 \sigma_i^2\right]\left[\sum_{i=1}^{3}\left(B_i'\left(\tau_2\right)\right)^2 \sigma_i^2\right]}},$$

where factors volatilities σ_i come from the the SDE $df_i\left(t\right) = \sigma_i dw_i\left(t\right)$ and

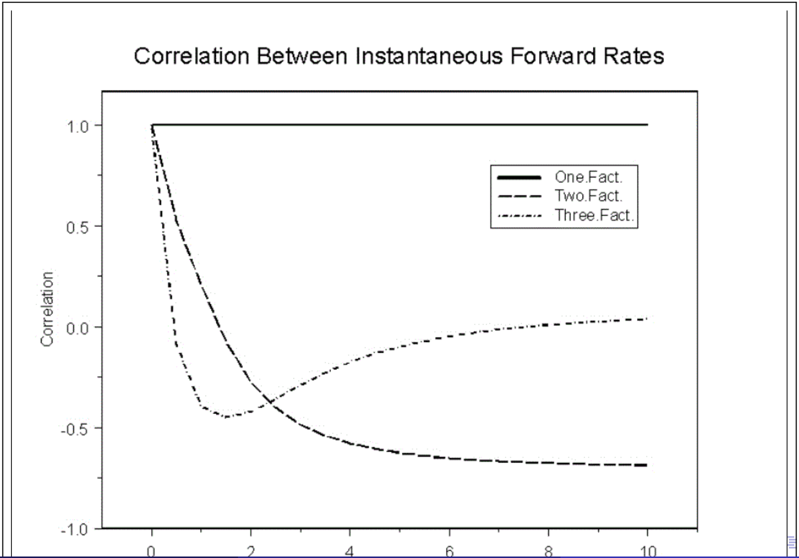
$$\begin{aligned} B_i'\left(\tau\right) &= \frac{d}{d\tau} B_i\left(\tau\right) \\ &= \frac{d}{d\tau} l_i\left(\tau\right)\tau = \sum_{j=1}^{m_i} b_{ij} \exp\left(-a_{ij}\tau\right). \end{aligned}$$

Price for Reduced Dimension

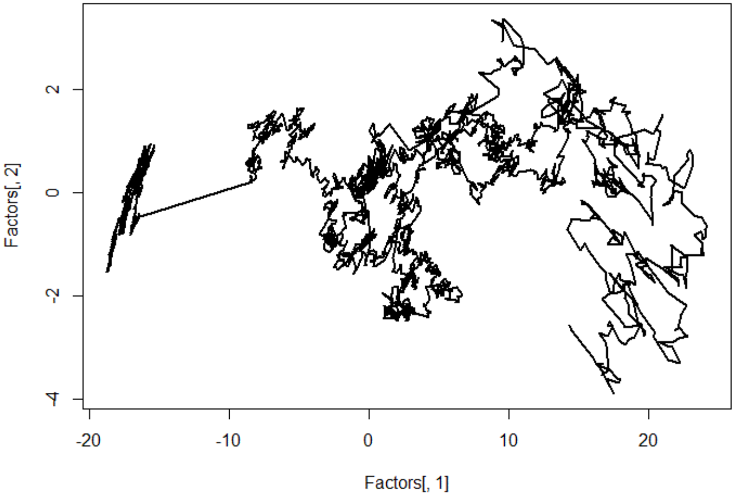
- The following slide shows the structure of correlations between instantaneous forward rates generated by one-, two- and three-factor statistical model: $\rho\left(\Delta F\left(t,t\right),\Delta F\left(t,t+\tau\right)\right)$, $\tau > 0$.
- Compare these correlations with the discussion in [1, p.37, p. 67-72].
- Correlation between forward rates decays growing τ.
- The rate of this decorrelation is an important market characteristic.
- There is a trade-off between model simplicity expressed in number of factors and ability to capture realistic correlation structure.
- One-factor model can reproduce any volatility term structure of the forward rates if we make the factor loading flexible enough. But it implies perfect correlations between forward rates.
- Any model with few driving factors has restricted ability to explain correlations.

Riccardo Rebonato, 1998, Interest-Rate Option Models: Understanding, Analyzing and Using Models for Exotic Interest-Rate Options, John Wiley & Sons.

Correlation Structure Between Instantaneous Forward Rates



Explore History through Factors



Concept of VaR I

- The industry standard risk measure approved by the regulators is VaR.
- If today (t) portfolio value is $\Pi(t)$, how low it can be at time horizon $(t+\Delta)$?
  Let $q_{0.01}$ be 1%-quantile of the distribution of $\Delta\Pi(t)$ over the period of Δ:

  $$q_{0.01} = \{x : \mathbf{P}\{\Delta\Pi(t) < x\} \leq 0.01, \mathbf{P}\{\Delta\Pi(t) < x + 0\} \geq 0.01\}$$

  The value $q_{0.01}$ is called *value at risk* or 99% VaR.
- VaR suggests that all losses that occur with the cumulative probability less than 0.01 can be ignored. Then the risk is measured as the worst outcome that cannot be ignored.

Concept of VaR II

- Often the value of portfolio depends on some state variable ξ and we know only the distribution of ξ, but not of the portfolio.
- Let the change of the portfolio during the time interval $(t, t+\Delta)$ be $F(\xi) = \Delta\Pi(t)$. Then a common way of finding VaR is $VaR = F\left(q_{0.01,\xi}\right)$ where is the quantile of ξ.
- This method is based on the assumption that the inverse function $F^{-1}\left(q_{0.01,F}\right)$ is known and

$$\mathbb{P}\left\{F(\xi) < q_{0.01,F}\right\} = \mathbb{P}\left\{\xi < F^{-1}\left(q_{0.01,F}\right) = q_{0.01,\xi}\right\}.$$

- Following this approach we can select factors of the statistical model as state variables.
- For simplicity we use the statistical model with only 2 factors. This model suggests that the values of all instruments at time $t+\Delta$ are defined by the factor volatilities and correlations.

Concentration Ellipse and VaR on the Plane I

Let $df_1 = y, df_2 = x$ be the daily changes of the factors.
Assume that the pair $\{x, y\}$ has joint normal distribution with $\mathbb{E}[x] = \mathbb{E}[y] = 0, \mathbb{V}[y] = \sigma_1^2, \mathbb{V}[x] = \sigma_2^2$, covariance $\sigma_{1,2}$ and correlation $\rho = \frac{\sigma_{1,2}}{\sigma_1 \sigma_2}$.
Then define the *ellipse of equal probabilities* or *concentration ellipse* as

$$\Phi(\lambda) = \left\{ x, y : \frac{1}{2(1-\rho^2)} \left(\frac{y^2}{\sigma_1^2} - 2\rho \frac{xy}{\sigma_1 \sigma_2} + \frac{x^2}{\sigma_2^2} \right) = \lambda^2 \right\}.$$

For such ellipse

$$\mathbb{P}\{\{x, y\} \in \Phi(\lambda)\} = 1 - \exp\left(-\lambda^2\right).$$

Standard concentration ellips, $\lambda = \sqrt{2}$, is especially convenient: the first and the second order moments of the 2-dimensional normal distribution are equal to the first and the second order moments of the uniform distribution on the ellipse.

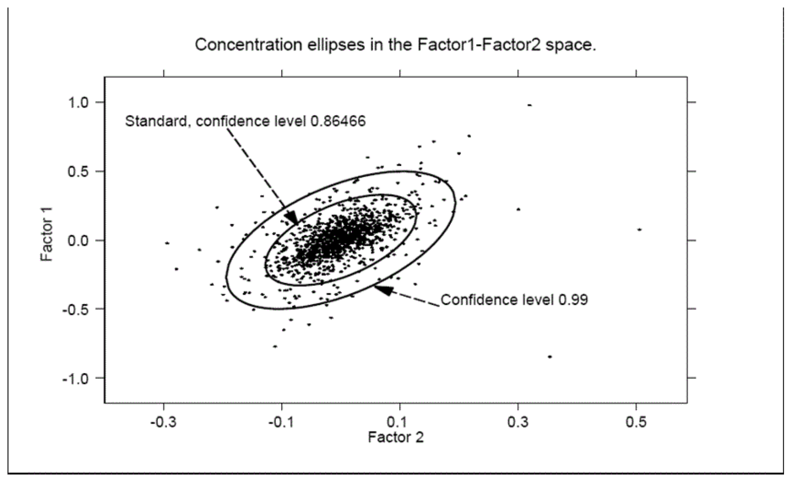
Concentration Ellipse and VaR on the Plane II

For the standard concentration ellipse the confidence level is

$$\mathbb{P}\left\{\{x,y\}\in\Phi\left(\sqrt{2}\right)\right\}=1-\exp\left(-2\right)=0.86466.$$

Such confidence level is lower than the industry requirement.
To make the confidence level of the ellipse 99% solve the equation $1-\exp\left(-\lambda^2\right)=.99$ and find $\lambda=\sqrt{4.6052}$.
Then the scenario method of risk calculation suggests calculating the change in the portfolio value at some points on the ellipse and choose the worst case.
We can use historical sample of factors to calculate the joint distribution of the factor changes.
The following slide shows the scatterplot of the factor changes, the standard concentration ellipse and the ellipse with the confidence level of 99%.

Concentration Ellipses



Rotation of Gaussian Vectors I

It is convenient to have the factor changes uncorrelated.
Any vector with normal distribution can be rotated by some orthogonal transformation, to make the components of the vector uncorrelated.
The factors at time t are $f(t) = \{f_1(t), f_2(t)\}$and the daily increments are $\Delta f(t) = \{\Delta f_1(t), \Delta f_2(t)\}$with the covariance matrix

$$\Sigma(t) = \{\sigma_{i,j}(t); i, j = 1, 2\} = \mathbb{E}\left[\Delta f(t)^T \Delta f(t)\right].$$

Assume that this covariance matrix is known at time t and remains constant at least during the period $[t, t + \Delta]$.
Let $C(t)$ be such matrix that

$$\Sigma(t) = C(t)\Lambda C^T(t).$$

It rotates the factors $f(t)$ to $h(t) = \{h_1(t), h_2(t)\} = f(t)C(t)$.

Rotation of Gaussian Vectors II

The argument t means that $C(t)$ is known at t and remains constant at least as long as $\Sigma(t)$ remains constant. Then

$$\Delta h(t) = \Delta f(t)\, C(t)$$

and the components of $\Delta h(t)$ are uncorrelated

$$\begin{aligned} \mathbb{E}\left[\Delta h(t)^T \Delta h(t)\right] &= \mathbb{E}\left[C^T(t)\, \Delta f(t)^T \Delta f(t)\, C(t)\right] \\ &= C^T(t)\, \Sigma(t)\, C(t) = \Lambda \end{aligned}$$

with the variances equal to the eigenvalues of the covariance matrix $\Sigma(t)$.

Factor Rotation I

The formula for rates in the statistical model is $Y_0(t,\tau) = f(t)\,L^T(\tau)$ or

$$\begin{aligned}
Y(t,\tau) - L_0(\tau) &= \sum_{i=1}^{2} l_i(\tau)\, f_i(t)\,, \\
L(\tau) &= \{l_j(\tau_i)\,; i = 1,\ldots,n; j = 1,2\} \\
&= \begin{bmatrix} l_1(\tau_1) & l_2(\tau_1) \\ \vdots & \vdots \\ l_1(\tau_n) & l_2(\tau_n) \end{bmatrix}.
\end{aligned}$$

Apply $C(t)$ to $f(t)$:

$$Y_0(t,\tau) = f(t)\,L^T(\tau) = f(t)\,C^T(t)\,C(t)\,L^T(\tau) = h(t)\,\tilde{L}^T(t,\tau)\,,$$

Factor Rotation II

$$
\begin{aligned}
\tilde{L}(t,\tau) &= L(t,\tau)\,C^T(t) \\
&= \begin{bmatrix} l_1(\tau_1) & l_2(\tau_1) \\ \vdots & \vdots \\ l_1(\tau_n) & l_2(\tau_n) \end{bmatrix} \begin{bmatrix} c_{1,1}(t) & c_{2,1}(t) \\ c_{1,2}(t) & c_{2,2}(t) \end{bmatrix} \\
&= \{\tilde{l}_j(t,\tau_i) = l_1(\tau_i)\,c_{j,1}(t) + l_2(\tau_i)\,c_{j,2}(t); \\
&\quad i = 1,\ldots,n; j = 1,2\}.
\end{aligned}
$$

Note again that the argument t in $\tilde{l}(t,\tau)$ means that this matrix is known at t and remains constant as long as the correlation structure of the factor increments $\Delta f(t)$ remains constant, but at least until $t+\Delta$.
The new loadings $\tilde{L}(t,\tau)$ eliminate the correlation structure between the factor increments.

Factor Rotation III

Since

$$l_i(\tau) = \sum_{j=1}^{m_i} b_{ij} \frac{1 - \exp(-a_{ij}\tau)}{a_{ij}\tau}$$

the new loadings are

$$\begin{aligned} \tilde{l}_i(t,\tau) &= l_1(\tau)\, c_{i,1}(t) + l_2(\tau)\, c_{i,2}(t) \\ &= \sum_{j=1}^{m_1} b_{1j} c_{i,1}(t) \frac{1 - \exp(-a_{1j}\tau)}{a_{1j}\tau} \\ &\quad + \sum_{j=1}^{m_2} b_{2j} c_{i,2}(t) \frac{1 - \exp(-a_{2j}\tau)}{a_{2j}\tau}. \end{aligned}$$

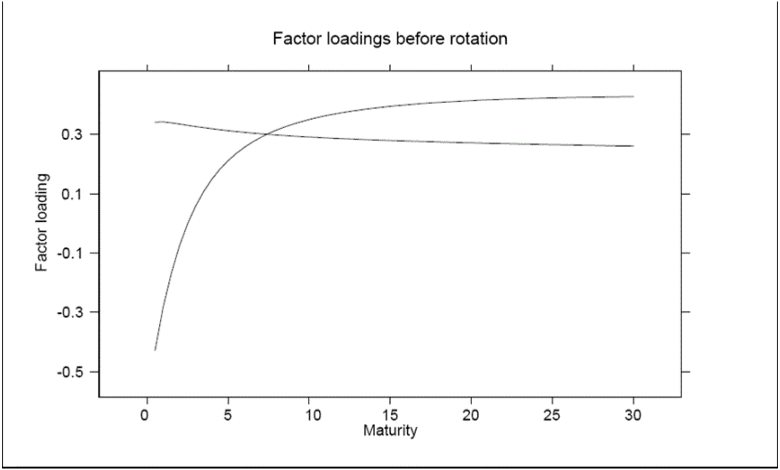
If the factor loading $l_i(\tau)$ is defined by the vector $\{a_i, b_i\} = \{a_{i,1}, \ldots, a_{i,m_i}, b_{i,1}, \ldots, b_{i,m_i}\}$ then the new factor loading $\tilde{l}_i(t,\tau)$ is defined by the vector $\{\tilde{a}_i, \tilde{b}_i\} = \{a_1, a_2, c_{i,1}(t)\, b_1, c_{i,2}(t)\, b_2\}$.

Loadings and Volatilities Before and After Rotation I

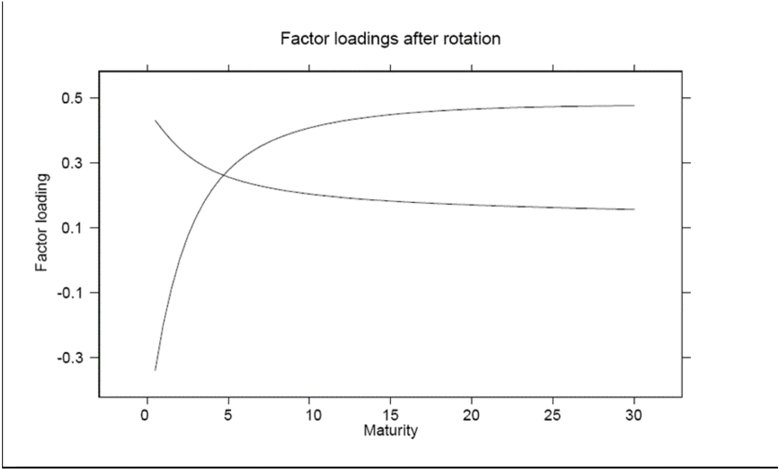
With the parameters a and b of the factor loadings that we fit to the PCA loadings we can see the results of rotation on the next two slides.
Rotation does not change the mathematical expectations of the factor increments.
The following table shows the variances, covariance coefficient and the correlation coefficient of the factor changes before and after the rotation.
The rotation made the correlation matrix of daily factors changes diagonal.

	Original	Rotated
Factor1 variance	0.027	0.02862
Factor2 variance	0.0047	0.002813
Cross-covariance	0.0057	$7e - 018$
Cross-correlation	0.5405	$8e - 016$

Parametric Factor Loadings Before Rotation



Parametric Factor Loadings After Rotation



Concentration Ellipse After Rotation

After the rotation the concentration ellipse has the form

$$\frac{y^2}{\tilde{\sigma}_1^2} + \frac{x^2}{\tilde{\sigma}_2^2} = 2\lambda^2,$$

where $\tilde{\sigma}_1^2$ and $\tilde{\sigma}_2^2$ are new variances, equal to the eigenvalues of the initial covariance matrix.

Note that the actual observations that fall outside the rotated concentration ellipse are much more frequent than the theory based on normal distribution predicts.

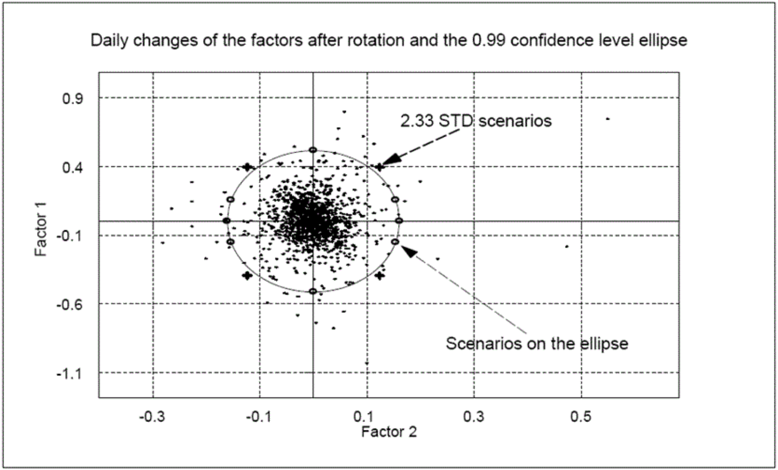
It is 45 out of 1250 observations or 3.6% of the sample.

What is even worse, the outliers are not distributed uniformly in time.

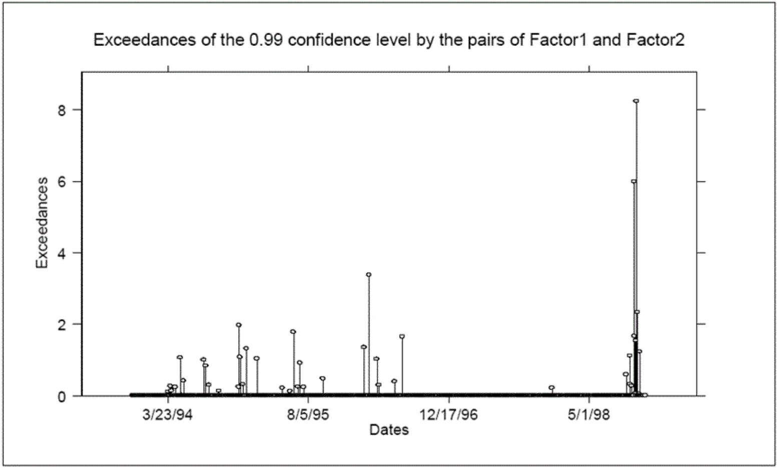
Look at slide showing the point process of the exceedances.

The normal distribution assumption and VaR as measure of risk may cause big unpredicted losses.

Concentration Ellipse After Rotation



Outside the Ellipse



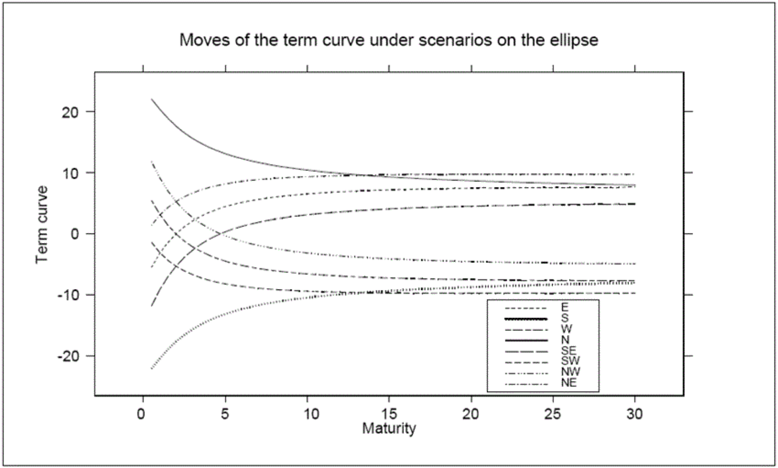
Scenarios on the Ellipse

The 99%-ellipse after the rotation shows the set of scenarios that we choose for the analysis:

N	$y=\sqrt{2}\lambda\tilde{\sigma}_1$	$x=0$
W	$y=0$	$x=-\sqrt{2}\lambda\tilde{\sigma}_2$
S	$y=-\sqrt{2}\lambda\tilde{\sigma}_1$	$x=0$
E	$y=0$	$x=\sqrt{2}\lambda\tilde{\sigma}_2$
NW	$y=\lambda\tilde{\sigma}_1\tilde{\sigma}_2\frac{\sqrt{2}}{\sqrt{\tilde{\sigma}_2^2+\tilde{\sigma}_1^2}}$	$x=-\lambda\tilde{\sigma}_1\tilde{\sigma}_2\frac{\sqrt{2}}{\sqrt{\tilde{\sigma}_2^2+\tilde{\sigma}_1^2}}$
SW	$y=-\lambda\tilde{\sigma}_1\tilde{\sigma}_2\frac{\sqrt{2}}{\sqrt{\tilde{\sigma}_2^2+\tilde{\sigma}_1^2}}$	$x=-\lambda\tilde{\sigma}_1\tilde{\sigma}_2\frac{\sqrt{2}}{\sqrt{\tilde{\sigma}_2^2+\tilde{\sigma}_1^2}}$
SE	$y=-\lambda\tilde{\sigma}_1\tilde{\sigma}_2\frac{\sqrt{2}}{\sqrt{\tilde{\sigma}_2^2+\tilde{\sigma}_1^2}}$	$x=\lambda\tilde{\sigma}_1\tilde{\sigma}_2\frac{\sqrt{2}}{\sqrt{\tilde{\sigma}_2^2+\tilde{\sigma}_1^2}}$
NE	$y=\lambda\tilde{\sigma}_1\tilde{\sigma}_2\frac{\sqrt{2}}{\sqrt{\tilde{\sigma}_2^2+\tilde{\sigma}_1^2}}$	$x=\lambda\tilde{\sigma}_1\tilde{\sigma}_2\frac{\sqrt{2}}{\sqrt{\tilde{\sigma}_2^2+\tilde{\sigma}_1^2}}$

Using the factor loadings after the rotation we can calculate the changes of the term curve under the scenarios on the ellipse (see next slide).

Term Curve Scenarios on the Ellipse



Scenarios Based on Standard Deviations

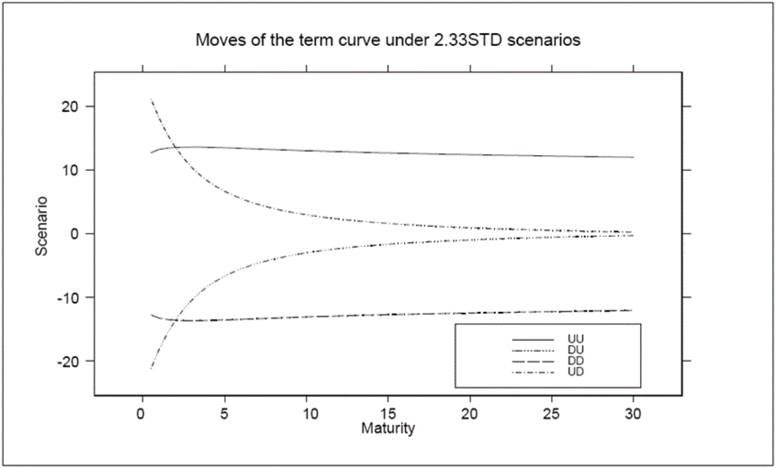
In one-factor case there is a very simple way of VaR calculation: estimate the standard deviation and make VaR equal to the value of 2.33σ.
This method is often used in the multidimensional case. For the 2-factor space the scenarios based on standard deviations are

UU	$y = 2.33\tilde{\sigma}_1$	$x = 2.33\tilde{\sigma}_2$
UD	$y = 2.33\tilde{\sigma}_1$	$x = -2.33\tilde{\sigma}_2$
DU	$y = -2.33\tilde{\sigma}_1$	$x = 2.33\tilde{\sigma}_2$
DD	$y = -2.33\tilde{\sigma}_1$	$x = -2.33\tilde{\sigma}_2$

These scenarios are outside the 99% ellipse, they are more conservative.
The corresponding changes of the term curve are shown on Following slide.
Both sets of scenarios have the same flow: we do not search inside the ellipse. See [1] for one solution of the problem and for more information on the scenario-based approach to the risk calculation.

Farshid Jamshidian, Yu. Zhu, 1997, Scenario Simulation: Theory and Methodology, Finance and Stochastics,1, 43-67.

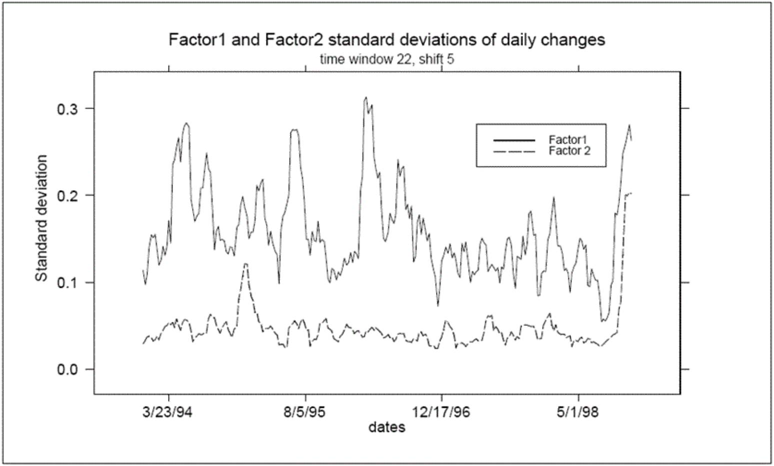
Term Curve Scenarios Based on Standard Deviations



Scenarios for volatilities

- We know that scenarios for the factors can only define the change in the value of the simplest instruments, like bonds and vanilla swaps. In most of the cases we also need scenarios for the factor volatilities.
- As a first approximation we can follow the same approach to scenarios definition.
- Build the time series of standard deviations of the factor increments using some time window and shift . The next slide shows the time series of historical volatilities of the factors.
- Then build the 99% confidence level ellipse for the daily changes of the standard deviations.
- Now choose points on this ellipse as scenarios for the VaR calculation.
- Note that the distribution of the changes of the standard deviations is even less normal than in the case of the factor changes.

Time Series of Historical Volatilities



Concentration Ellipse for Volatilities