Principal components of stock market dynamics

Methodology and applications in brief (to be updated…)

Andrei Bouzaev, bouzaev@ya.ru
Why principal components are needed

Objectives

• understand the evidence of more than one factor driving the stock market, market factor should be supplemented with others, possibly less significant

• identify set of factor components driving stock market and its basic segments

• set up multi-factor model that is better than one-factor for monitoring of market dynamics, hedge, pricing, risk-valuation
Similar to modeling the shape of the yield curve

*Modelling the shape of the yield curve – very similar process*

- evidence of more than one factor driving the term structure
- principal components of dynamics are needed to hedge the shape of the yield curve in practice

**Yield curve as at 9th February 2005 for USD**

**The following three factors are used: level, slope, curvature**

**Sensitivities of zero-coupon yields to the three factors**
In reality, do we have only one factor?

Correlation matrix indicates that difference in return dynamics is increasing by the difference in market equity.

<table>
<thead>
<tr>
<th></th>
<th>LO10</th>
<th>DEC2</th>
<th>DEC3</th>
<th>DEC4</th>
<th>DEC5</th>
<th>DEC6</th>
<th>DEC7</th>
<th>DEC8</th>
<th>DEC9</th>
<th>HI10</th>
</tr>
</thead>
<tbody>
<tr>
<td>LO10</td>
<td>1,00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DEC2</td>
<td>1,00</td>
<td>0,96</td>
<td>0,93</td>
<td>0,91</td>
<td>0,88</td>
<td>0,86</td>
<td>0,85</td>
<td>0,81</td>
<td>0,79</td>
<td>0,70</td>
</tr>
<tr>
<td>DEC3</td>
<td>1,00</td>
<td>0,97</td>
<td>0,97</td>
<td>0,95</td>
<td>0,92</td>
<td>0,91</td>
<td>0,88</td>
<td>0,86</td>
<td>0,82</td>
<td>0,78</td>
</tr>
<tr>
<td>DEC4</td>
<td>1,00</td>
<td>0,98</td>
<td>0,97</td>
<td>0,96</td>
<td>0,94</td>
<td>0,92</td>
<td>0,92</td>
<td>0,92</td>
<td>0,84</td>
<td>0,82</td>
</tr>
<tr>
<td>DEC5</td>
<td>1,00</td>
<td>0,98</td>
<td>0,97</td>
<td>0,96</td>
<td>0,94</td>
<td>0,92</td>
<td>0,92</td>
<td>0,92</td>
<td>0,84</td>
<td>0,82</td>
</tr>
<tr>
<td>DEC6</td>
<td>1,00</td>
<td>0,98</td>
<td>0,97</td>
<td>0,96</td>
<td>0,96</td>
<td>0,96</td>
<td>0,96</td>
<td>0,96</td>
<td>0,87</td>
<td>0,82</td>
</tr>
<tr>
<td>DEC7</td>
<td>1,00</td>
<td>0,98</td>
<td>0,97</td>
<td>0,96</td>
<td>0,97</td>
<td>0,97</td>
<td>0,97</td>
<td>0,97</td>
<td>0,91</td>
<td>0,91</td>
</tr>
<tr>
<td>DEC8</td>
<td>1,00</td>
<td>0,98</td>
<td>0,97</td>
<td>0,96</td>
<td>0,98</td>
<td>0,98</td>
<td>0,98</td>
<td>0,98</td>
<td>0,93</td>
<td>0,93</td>
</tr>
<tr>
<td>DEC9</td>
<td>1,00</td>
<td>0,98</td>
<td>0,97</td>
<td>0,96</td>
<td>0,98</td>
<td>0,98</td>
<td>0,98</td>
<td>0,98</td>
<td>0,95</td>
<td>0,95</td>
</tr>
</tbody>
</table>

• Correlation matrix indicates that difference in return dynamics is increasing by the difference in market equity.

• Are there underlying factors that influence different portfolios in different manner?
In reality, do we have only one factor?

Correlation between Russia stock market portfolios formed on size. Data: MICEX capitalisation indices, based on daily price levels (left) and daily returns (right), 2005-2011

<table>
<thead>
<tr>
<th></th>
<th>BASE</th>
<th>STANDARD</th>
<th>HIGH</th>
</tr>
</thead>
<tbody>
<tr>
<td>BASE</td>
<td>1</td>
<td>0.97</td>
<td>0.92</td>
</tr>
<tr>
<td>STANDARD</td>
<td>1.00</td>
<td>0.96</td>
<td></td>
</tr>
<tr>
<td>HIGH</td>
<td>1.00</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- Correlation matrix indicates that difference in price dynamics is increasing by the difference in market equity.
- Are there underlying factors that influence different portfolios in different manner?
Identifying principal components*

- Let us find 3 factors that explain the dynamics of the ME portfolios
  - The analysis for more than 3 factors is a straightforward extension

- Assume the following dynamics of portfolios returns

\[
\begin{align*}
  r_1(t) &= \alpha_1 + \beta_{11}\Delta\phi_1(t) + \beta_{12}\Delta\phi_2(t) + \beta_{13}\Delta\phi_3(t) + \varepsilon_1(t) \\
  r_2(t) &= \alpha_2 + \beta_{21}\Delta\phi_1(t) + \beta_{22}\Delta\phi_2(t) + \beta_{23}\Delta\phi_3(t) + \varepsilon_2(t) \\
  &\vdots \\
  r_n(t) &= \alpha_n + \beta_{n1}\Delta\phi_1(t) + \beta_{n2}\Delta\phi_2(t) + \beta_{n3}\Delta\phi_3(t) + \varepsilon_n(t)
\end{align*}
\]

meaning that each factor impacts all of the returns with sensitivities \( \beta \) (alternatively the same regression for price levels can be used).

*The approach is adopted from modeling the shape of the yield curve, Fixed Income lecture notes at NES, professor Dmitri Makarov
Principal component analysis

• We look for factors that are implicit in the movements over time of the various returns

• To determine them we are going to apply Principal Component Analysis
  – Finding first principal component means that we are maximizing a weighted average of the R-square-s of the above regressions
  – After computing the first component, we can find the residuals, applying them instead of returns, and find the second factor, and so on
Principal components in practice (USA stock market)

- Using monthly returns for the US stock market between 1926 and 2000, we have the following sensitivities of size-based portfolios returns to the three factors:

<table>
<thead>
<tr>
<th>Portfolio</th>
<th>f1</th>
<th>f2</th>
<th>f3</th>
</tr>
</thead>
<tbody>
<tr>
<td>LO20</td>
<td>0.94</td>
<td>-0.314</td>
<td>0.12</td>
</tr>
<tr>
<td>QNT2</td>
<td>0.98</td>
<td>-0.151</td>
<td>-0.04</td>
</tr>
<tr>
<td>QNT3</td>
<td>0.99</td>
<td>-0.003</td>
<td>-0.09</td>
</tr>
<tr>
<td>QNT4</td>
<td>0.98</td>
<td>0.124</td>
<td>-0.09</td>
</tr>
<tr>
<td>HI20</td>
<td>0.93</td>
<td>0.350</td>
<td>0.12</td>
</tr>
</tbody>
</table>

Extraction Method: Principal Component Analysis.
Principal components in practice (Russia stock market)

Using daily price data for the Russia stock market between 2005 and 2011, we have the following sensitivities of size-based portfolios’ price level (1) and returns (2) to the three factors:

**Correlation matrix, price levels**

<table>
<thead>
<tr>
<th></th>
<th>(f_1)</th>
<th>(f_2)</th>
<th>(f_3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>BASE</td>
<td>0,98</td>
<td>-0,19</td>
<td>0,08</td>
</tr>
<tr>
<td>STANDARD</td>
<td>0,99</td>
<td>-0,05</td>
<td>-0,11</td>
</tr>
<tr>
<td>HIGH</td>
<td>0,97</td>
<td>0,24</td>
<td>0,04</td>
</tr>
</tbody>
</table>

**Correlation matrix, returns**

<table>
<thead>
<tr>
<th></th>
<th>(f_1)</th>
<th>(f_2)</th>
<th>(f_3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>D_BASE</td>
<td>0,90</td>
<td>0,41</td>
<td>0,12</td>
</tr>
<tr>
<td>D_STAND</td>
<td>0,96</td>
<td>-0,05</td>
<td>-0,29</td>
</tr>
<tr>
<td>D_HIGH</td>
<td>0,92</td>
<td>-0,35</td>
<td>0,18</td>
</tr>
</tbody>
</table>
Empirical components’ description

• First factor has roughly the same effect on all portfolio yields/prices. For this reason, it can be referred to as parallel shift/level. It explains most of the variation

• Second factor has negative impact on small stocks and positive impact on big stocks. This is a slope factor

• The third factor is positive at the small and big stocks, and negative at the medium stocks. It can be referred to as curvature
Returns or price levels?

- Applying factor analysis to returns or price levels should lead to the same results
  - Meaning “return factors” can be transformed mathematically into “level factors” correspondingly and vice versa
- Empirically there are some difficulties for return data
  - Positive correlation in levels with slightly lagged data can lead to negative correlation in returns
  - Daily returns are very “noisy”, for weekly or monthly returns a long period of data is needed
- For the above reasons price level components have been extracted for Russia stock market
Principal components in dynamics

- Three factor component in dynamic, Russia stock market (based on daily price levels)
Constructing “market curve”

Based on three factor components we can construct “stock market curve” which indicates underpriced/overpriced segments.

Based on Russia stock market data (daily price levels), 2005-2011.

Market curve, russian stock market - micex

F2, 5% +
F2, 5% -
F3, 5% +
F3, 5% -
11.11.2011
30.11.2011
30.12.2011

Normal shape of the curve: all market segments around their average

F2 slope factor

F3 curvature factor

F1 parallel shift

base
standard
high

Based on Russia stock market data (daily price levels), 2005-2011.
Application – multifactor asset pricing model (1)

• Expected return of any marketable security can be written as a function of the expected return of the efficient portfolio

\[ E[r_i] = r_f + \beta_i^{\text{eff}} \times (E[r_{\text{eff}}] - r_f) \]

• It is extremely difficult to identify portfolios that are efficient but we can use collection of portfolios from which the efficient portfolio can be constructed to measure risk

• Each factor (portfolio) captures different components of the systematic risk

• Multifactor models is significant improvement over the CAPM, it is widely used in academic literature and in practice to measure risk and to calculate cost of capital
Application – multifactor asset pricing model (2)

• The above described factors represent the following portfolios
  – The **level factor** is a market portfolio (to the extent it can be derived from stock market data)
  – The **slope factor** is small-minus-big portfolio. It is similar (but not identical) to SMB factor from Fama-French model).
  – The **curvature factor** is “small and big minus medium” portfolio.

• Taking into account the set of 3 factors the multifactor models becomes

\[ E[r_i] = r_f + \beta_i^{f1} (E[r_{f1}] - r_f) + \beta_i^{f2} (E[r_{f2}] - r_f) + \beta_i^{f3} (E[r_{f3}] - r_f) \]
Application – multifactor asset pricing model (3)

• Most popular multifactor model specification by now is based on market, small-minus-big and high-minus-low factors (see http://en.wikipedia.org/wiki/Fama%E2%80%93French_three-factor_model):

\[ E[r_i] = r_f + \beta_i^{Mkt} \times (E[r_{Mkt}] - r_f) + \beta_i^{SMB} \times (E[r_{SMB}] - r_f) + \beta_i^{HML} \times (E[r_{HML}] - r_f) \]

• But these factors were selected empirically whereas level, slope and curvature factors are derived mathematically (using principal components analysis)
Further development

• Factor components methodology and components’ values are published on http://fmlab.hse.ru/

• Ready for any discussion and cooperation with individuals/academics/financial institutions of further development and calculation of the presented factor components in form of stock indices or returns.

• Contacts: Andrei Bouzaev, bouzaev@ya.ru, (495) 9696 225