Modeling Euro Bonds via multivariate nonparametric regression

EUMOPTFIN 3, Bergamo, 18th May 2004

Martin Hanek
Charles University in Prague, Czech Republic
Department of Probability and Mathematical Statistics
Data

- UniCredit Banca Mobiliare Milan, already studied
- 729 files of risk-free Euro yields for maturities 1M, 6M, 9M, 1Y, 2Y, 3Y,...10Y from 15.3.2000 to 31.3.2003
- 2880 files with credit spreads of Eurobonds with maturities 2Y-30Y from 4.4.2000 to 26.3.2003, altogether 323,977 observations
- Euro Bonds: denominated in a currency other than the issuer’s, OTC (pricing?), ratings (AAA,AA,A,BBB)
<table>
<thead>
<tr>
<th></th>
<th>AAA</th>
<th>AA</th>
<th>A</th>
<th>BBB</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>min</strong></td>
<td>-49.79</td>
<td>-27.78</td>
<td>2.60</td>
<td>14.02</td>
</tr>
<tr>
<td><strong>1. quartile</strong></td>
<td>16.57</td>
<td>31.19</td>
<td>64.43</td>
<td>111.21</td>
</tr>
<tr>
<td><strong>median</strong></td>
<td>23.94</td>
<td>42.67</td>
<td>85.31</td>
<td>145.02</td>
</tr>
<tr>
<td><strong>mean</strong></td>
<td>25.99</td>
<td>46.09</td>
<td>93.79</td>
<td>181.97</td>
</tr>
<tr>
<td><strong>3. quartile</strong></td>
<td>34.22</td>
<td>58.81</td>
<td>111.18</td>
<td>210.18</td>
</tr>
<tr>
<td><strong>max</strong></td>
<td>142.63</td>
<td>256.69</td>
<td>900.25</td>
<td>1614.46</td>
</tr>
</tbody>
</table>

Table 1: Spreads (bps)
Zero rates

yield in %

maturity in days

time in days
spreads AAA

![3D scatter plot with axes labeled as:
- x-axis: time in days
- y-axis: maturity in days
- z-axis: spread in bps

The plot shows a 3D visualization of the spread values over time and maturity, with data points scattered across the axes.]
spreads A

![Graph of spreads A showing time in days, maturity in days, and spread in bps.](image-url)
Methods

Parametric methods do not work - not robust, inflexible, for practical purposes (prediction) do not need an economically justified model ⇒ SMOOTH

Data \((X_i, Y_i), i = 1 \ldots n, \exists m\) smooth

\[ Y_i = m(X_i) + \varepsilon_i, \quad E[\varepsilon_i] = 0, E[\varepsilon_i \varepsilon_j'] = \sigma^2 I_n, \quad i, j = 1, \ldots, n \quad (1) \]

Classical nonparametric regression estimate \(\hat{m}(x_0)\) at \(x_0\)

\[ \min_{\theta} \left[ \sum_{i=1}^{n} (Y_i - \hat{m}(X_i, \theta))^2 K_h(X_i - x_0) \right], \quad (2) \]
LOESS

=LOWESS (Locally Weighted Scatterplot Smoothing), Cleveland 1988 K-Nearest Neighbor local polynomial iterative nonparametric regression estimate (possibly multivariate)

K-NN: \((x_0 - d, x_0 + d)\), where \(d\) is the distance of K-th nearest neighbor from \(x_0\)

Tri-cubic weight function \(W(X_k) = \left(1 - \frac{|x_0-X_k|}{d}\right)^3\) for \(\frac{|x_0-X_k|}{d} \leq 1\)

Estimate \(\hat{m}(x_0) = \hat{a} + \hat{b}(x_0 - x_0) = \hat{a}\), where

\[(\hat{a}, \hat{b}) = \arg\min_{a,b \in \mathbb{R}} \sum_{k=1}^{n} W(X_k) \cdot (Y_k - a - b(X_k - x_0))^2\] \hspace{1cm} (3)
Robustification to guard against outliers ⇒ iterate ($< 5$ times)

Residuals $\hat{\varepsilon}_i = Y_i - \hat{m}(X_i) \Rightarrow m = \text{median}(|\hat{\varepsilon}_i|) \Rightarrow \hat{e}_i = \hat{\varepsilon}_i/(6\cdot m)$ (for $\varepsilon \sim N(0, \sigma^2)$)

$6m$ estimates $4\sigma$⇒

$$G(X_k) = \begin{cases} 
(1 - \hat{e}_k^2)^2 & \text{if } |\hat{e}_k| \leq 1 \\
0 & \text{otherwise} 
\end{cases} \quad (4)$$

Now combine $W(X_k)$ and $G(X_k)$ and estimate $\hat{m}_{iter+1}(x_0) = \hat{a}_{iter+1} + \hat{b}_{iter+1}(x_0 - x_0) = \hat{a}_{iter+1}$ kde

$$(\hat{a}_{iter+1}, \hat{b}_{iter+1}) = \arg \min_{a,b \in \mathbb{R}} \sum_{k=1}^{n} G_{iter}(X_k)W(X_k)(Y_k - a - b(X_k - x_0))^2 \quad (5)$$
Models

Symbolically $Yield = m(maturity, time) + \varepsilon$

Multivariate methods (Loess and Nonparametric regression) oversmooth, cannot include the daily variability. OK for a trend estimate but not for the daily prediction.

Add a regressor $\hat{Yield}_0 = \hat{m}(maturity_0, time_0)$ by univariate nonparametric regression and estimate $Spread_0 = m(maturity_0, time_0, \hat{Yield}_0) + \varepsilon$

<table>
<thead>
<tr>
<th>method</th>
<th>$f(\varepsilon)$</th>
<th>Min</th>
<th>1st Qu.</th>
<th>Median</th>
<th>Mean</th>
<th>3rd Qu.</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>loess</td>
<td>$\varepsilon$</td>
<td>-0.38</td>
<td>-0.03</td>
<td>-0.00</td>
<td>0.01</td>
<td>0.03</td>
<td>3.17</td>
</tr>
<tr>
<td>univ</td>
<td>$\varepsilon$</td>
<td>-0.43</td>
<td>-0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.32</td>
</tr>
<tr>
<td>univ</td>
<td>$</td>
<td>\varepsilon</td>
<td>$</td>
<td>0.00</td>
<td>0.001</td>
<td>0.003</td>
<td>0.02</td>
</tr>
<tr>
<td>univ</td>
<td>$\varepsilon^2$</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.002</td>
<td>0.00</td>
<td>0.19</td>
</tr>
</tbody>
</table>
Zerorates–nonparametric smooth $h=(5,0.7)$
Zerorates–loess smooth span=0.005
Zerorates–univariate nonparametric smooth $h=0.5$
Zerorates–residuals loess

![Graph showing zerorates residuals loess with axes for time in days, maturity in years, and residual value.](image)
Zerorates–residuals univariate nonparametric smooth

![Graph showing Zerorates–residuals univariate nonparametric smooth relationship with time in days, maturity in years, and residual.](image-url)
spreads AAA Loess (0.15)
spreads AA Loess (0.3)
spreads A Loess (0.3)
spreads BBB Loess (0.3)
<table>
<thead>
<tr>
<th>rating</th>
<th>SSE/n</th>
<th>1%</th>
<th>5%</th>
<th>10%</th>
<th>30%</th>
<th>50%</th>
<th>70%</th>
<th>90%</th>
<th>95%</th>
<th>99%</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAA</td>
<td>102.1</td>
<td>-35.1</td>
<td>-15.1</td>
<td>-10.4</td>
<td>-4.0</td>
<td>0.3</td>
<td>4.3</td>
<td>9.8</td>
<td>12.7</td>
<td>23.7</td>
</tr>
<tr>
<td>AA</td>
<td>305.2</td>
<td>-53.5</td>
<td>-30.6</td>
<td>-21.2</td>
<td>-6.4</td>
<td>0.8</td>
<td>7.9</td>
<td>20.0</td>
<td>25.5</td>
<td>35.0</td>
</tr>
<tr>
<td>A</td>
<td>2145.0</td>
<td>-165.4</td>
<td>-80.6</td>
<td>-51.4</td>
<td>-14.7</td>
<td>1.4</td>
<td>15.5</td>
<td>32.9</td>
<td>40.7</td>
<td>52.2</td>
</tr>
<tr>
<td>BBB</td>
<td>15461.1</td>
<td>-547.4</td>
<td>-245.6</td>
<td>-136.3</td>
<td>-34.9</td>
<td>4.6</td>
<td>30.4</td>
<td>62.8</td>
<td>75.7</td>
<td>99.2</td>
</tr>
</tbody>
</table>
Thank you for the attention

Question, comments,.....???