Estimating Market Conditions Based on Random Matrix Theory

MFM Modeling Workshop 2016
Final Report
Introduction

• The purpose of the project is to compare different indicators of financial crisis that have been proposed recently.

• Theory : Kornprobst and Douady’s paper, A Practical Approach to Financial Crisis Indicators Based on Random Matrices.

• Data : We are analyzing how the correlation of 226 assets in the S&P500 over a 25-year span change during times of crisis.
Empirical dataset

- In our empirical dataset, each row contains an asset, and each column represents an observation at a date \( t \).
- Calculate the log-return of each asset.
- Construct rolling window \( T \) of 150 days in the past.
- Get covariance matrix \( CV(t) \) between each asset at each date \( t \) using the data of past 150 days.

\[
ROL_t = ROL_t(j,:) - \text{mean}(ROL_t(j,:))
\]

\[
CV_t = \frac{1}{T} \times ROL_t \times (ROL_t)'
\]

- They compared the eigenvalue of covariance matrix with reference distribution by taking the Hellinger Distance.
3 indicators for financial crisis

• Compare eigenvalue of covariance matrix of empirical dataset with reference distribution
• The further away the empirical distribution drifts from the reference, the more likely the market is about to experience a crisis

• Three reference distribution:
  • 1. Marchenko Pastur distribution
  • 2. Standard Normal distribution
  • 3. Student T distribution
Marchenko Pastur density

\[ f(x) = \frac{1}{2\pi \sigma^2 \gamma} \sqrt{\left( \lambda^+ - x \right) \left( x - \lambda^- \right)} \quad \lambda^- < x < \lambda^+ \]

\[ \lambda^- = \sigma^2 \left( 1 + \sqrt{\gamma} \right)^2 \]

\[ \lambda^- = \sigma^2 \left( 1 - \sqrt{\gamma} \right)^2 \]

\[ \gamma = \frac{l}{c} \]
The Hellinger Distance

• It is used to evaluate the difference between empirical distribution and reference distribution at each point.

• Formula: \( D = \left( \sum_{i=1}^{n} \sqrt{P(X_i)} - \sqrt{Q(X_i)} \right) \)

• For the MP and correlated Gaussian distribution, the bigger the Hellinger distance, the more likely a crisis would occur.

• For student t (fat tails), the smaller the Hellinger Distance indicated an increase in a crisis occurring.

Results of all three indicators
Criticisms

• The indicators focus on the eigenvalues within the reference distributions
• We focused on finding indicators that were based on the larger eigenvalues
Indicators

• First step: eigenvalue of covariance/correlation matrix

• Paper compares eigenvalue of covariance matrix of empirical dataset with reference distribution

• Use Marchenko Pastur distribution, normal distribution and student t distribution as reference

• focus on the noise part of eigenvalue: small elements of eigenvalue
Indicators

• We focus on the larger element of eigenvalue: the information part
• Calculate the **cut point** of eigenvalue:
• Comparing \( \lambda^+ = \sigma^2 \left( 1 + \sqrt{\gamma} \right)^2 \) with lambda in eigenvalue matrix

• \( \lambda_{\text{max}} \), proportion of informative eigenvalue, average correlation
Moving average

• Consider the average value of the last 200/50 days of the indicator at anytime.

• We make our new indicator =

  (average of 50 days) – (average of 200 days)

• Check the momentum on the indicator. See how the indicator is moving before crisis.
Red: Entropy1
Green: Entropy2/10
Blue: fraction_cnst_var
Blue-green: fraction_var
Purplish red: lambda_max
Yellow: mean_rho
Black: MixedRank/100
Dark Blue: rank/50
Orange: sum_lambda*5
Indicators from meaningful part:

We have currently looked at the following indicators:

1. Maximum value of eigen value: $\lambda_{max}$
2. Total variance of meaningful part: $\sum \lambda$
3. % of variance by the meaningful part: $\sum \lambda / \sigma^2$
4. Rank of the matrix where we consider the meaningful part
5. Mixed rank
6. Some estimation of entropy
7. Mean correlation: $<\rho>$
The observation for all the indicators are very similar so we focus on two indicators: $\lambda_{\text{max}}$ and $<\rho>$
We then look at the value of the indicators vs the log of daily returns.

As observed the value of the indicators don’t give us meaningful information!

What can we do next?
- look at the difference in the indicators and compare with the daily returns – Indicator on an indicator!
Interestingly this data tells use something – how change/jump in $\lambda_{\text{max}}$ is correlated with the returns!

Similar observation with change in $\langle \rho \rangle$ but with more fat (smeared)!

What can we do next?
- select at cut off and see if that gives us any hint!
The cut off was chosen manually to see meaningful result – may be more work is needed to figure out a sensible way to figure out a cut off

What can we do next?
- Quantify!
Not perfect but we do see that with a time window of 1 month or more our indicator give us better results
Conclusion

- Looked at the meaningful part of the covariance matrix instead of random part.
- Focused our at the following indicators:
  - Maximum value of eigen value: $\lambda_{\text{max}}$
  - Mean correlation: $\langle \rho \rangle$
- Looked at the daily jumps in the indicators which did give us some meaningful result

Future work:
- Look at average weekly or monthly jumps
- Consider different lengths of rolling window
- Explore more on the smile curve – whether the relationship would be informative?
Extra slides: