How to make loads of money using persistent homology

Jake Saunders

University of Sheffield

You can find these slides at `jses.site/talks`. This is not my own work.

Anticipating Market Crashes

- Label stocks in DJIA as $1, \dots, n$.

Anticipating Market Crashes

- Label stocks in DJIA as $1, \ldots, n$.
- $S_i(t)$ is price of stock i at end of day t.

Anticipating Market Crashes

- Label stocks in DJIA as $1, \ldots, n$.
- $S_i(t)$ is price of stock i at end of day t.
- $x_i(t) := \frac{S_i(t+1) - S_i(t)}{S_i(t)}$

Anticipating Market Crashes

- Label stocks in DJIA as $1, \ldots, n$.
- $S_i(t)$ is price of stock i at end of day t.
- $x_i(t) := \frac{S_i(t+1) - S_i(t)}{S_i(t)}$

Definition (Correlation Graph for Day t)

Weighted graph where

- the vertices are stocks,
- the weight of the edge between two stocks is $\sqrt{2(1 - c_{i,j}(t))}$.

Anticipating Market Crashes

- Label stocks in DJIA as $1, \ldots, n$.
- $S_i(t)$ is price of stock i at end of day t.
- $x_i(t) := \frac{S_i(t+1) - S_i(t)}{S_i(t)}$

Definition (Correlation Graph for Day t)

Weighted graph where

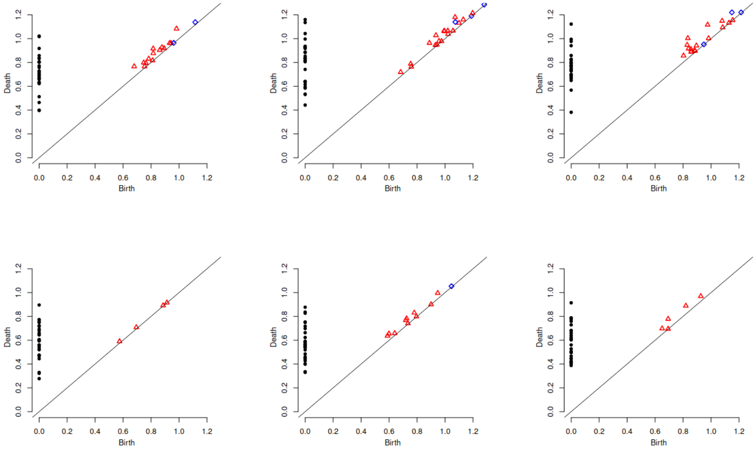
- the vertices are stocks,
- the weight of the edge between two stocks is $\sqrt{2(1 - c_{i,j}(t))}$.

Nasty Formula (for Pearson correlation coefficient)

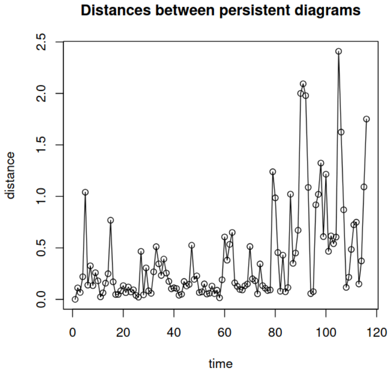
For fixed T (author uses $T = 15$),

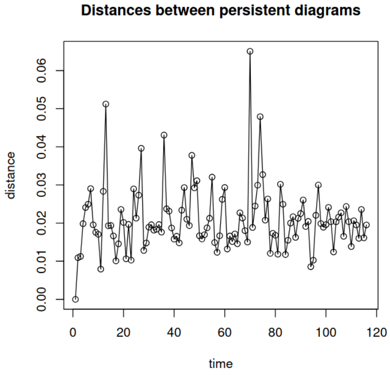
$$c_{i,j}(t) := \frac{\sum_{\tau=t-T}^{t} (x_i(\tau) - \overline{x_i})(x_j(\tau) - \overline{x_j})}{\sqrt{\sum_{\tau=t-T}^{t} (x_i(\tau) - \overline{x_i})^2} \sqrt{\sum_{\tau=t-T}^{t} (x_j(\tau) - \overline{x_j})^2}}.$$

Some Graphs

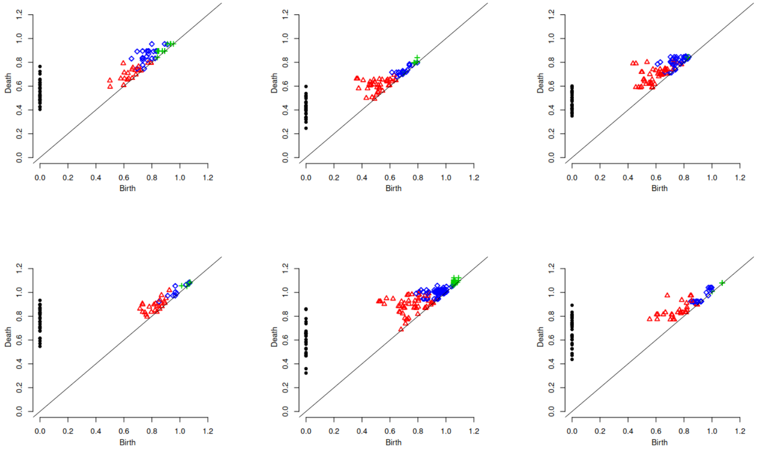


Some More Graphs

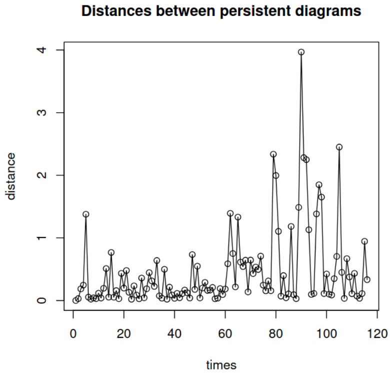


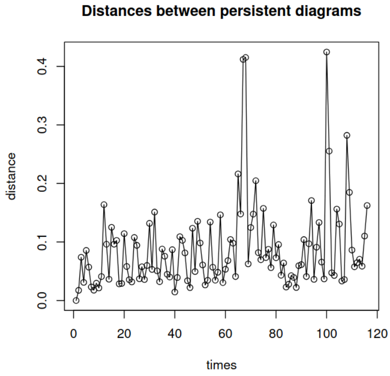


Further Graphs



Yet More Graphs





Flash Crash

- In 2010, the major US stock markets dropped more than 6 percent,

Flash Crash

- In 2010, the major US stock markets dropped more than 6 percent, then rebounded,

Flash Crash

- In 2010, the major US stock markets dropped more than 6 percent, then rebounded, in the space of 36 minutes.

Flash Crash

- In 2010, the major US stock markets dropped more than 6 percent, then rebounded, in the space of 36 minutes.
- Can this be predicted?

Flash Crash

- In 2010, the major US stock markets dropped more than 6 percent, then rebounded, in the space of 36 minutes.
- Can this be predicted?
- Futures – buy and agree to sell a stock at a fixed time in the future.

Flash Crash

- In 2010, the major US stock markets dropped more than 6 percent, then rebounded, in the space of 36 minutes.
- Can this be predicted?
- Futures – buy and agree to sell a stock at a fixed time in the future.
- Authors define an “abnormality index” based on the price of futures.

Flash Crash

- At minute j, let $x_j \in \mathbb{R}^3$ have coordinates corresponding to price of futures for S&P500, DJIA and NASDAQ.

Flash Crash

- At minute j, let $x_j \in \mathbb{R}^3$ have coordinates corresponding to price of futures for S&P500, DJIA and NASDAQ.
- Fix a time t, and consider the dataset consisting of $x_{t-50}, \ldots, x_t$.

Flash Crash

- At minute j, let $x_j \in \mathbb{R}^3$ have coordinates corresponding to price of futures for S&P500, DJIA and NASDAQ.
- Fix a time t, and consider the dataset consisting of $x_{t-50}, \ldots, x_t$.
- Compute norm of persistence diagram. Consider the 'exponential moving average/variance' of this quantity.

Flash Crash

- At minute j, let $x_j \in \mathbb{R}^3$ have coordinates corresponding to price of futures for S&P500, DJIA and NASDAQ.
- Fix a time t, and consider the dataset consisting of $x_{t-50}, \dots, x_t$.
- Compute norm of persistence diagram. Consider the 'exponential moving average/variance' of this quantity.
- Let EMA_1 be the first norm, $\mathrm{EMVar}_1 = 0$.

Flash Crash

- At minute j, let $x_j \in \mathbb{R}^3$ have coordinates corresponding to price of futures for S&P500, DJIA and NASDAQ.
- Fix a time t, and consider the dataset consisting of $x_{t-50}, \dots, x_t$.
- Compute norm of persistence diagram. Consider the 'exponential moving average/variance' of this quantity.
- Let EMA_1 be the first norm, $\text{EMVar}_1 = 0$.
- Define inductively

$$\begin{aligned}
\delta_i &:= i\text{th Norm} - \text{EMA}_{i-1}, \\
\text{EMA}_i &:= \text{EMA}_{i-1} + \alpha \cdot \delta_i, \\
\text{EMVar}_i &:= (1-\alpha)(\text{EMVar}_{i-1} + \alpha + \delta_i^2).
\end{aligned}$$

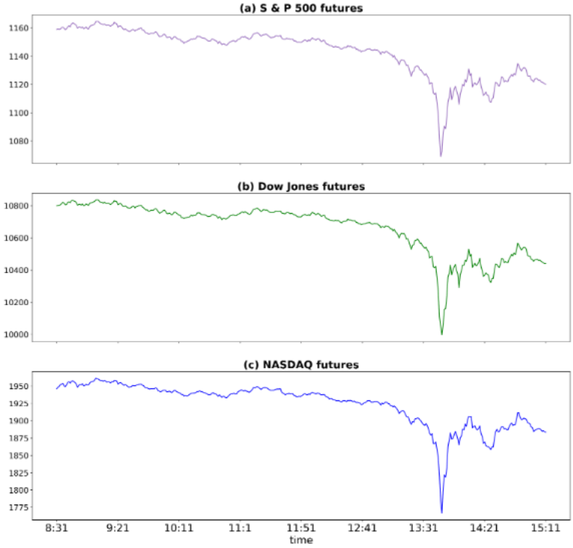
Flash Crash

- At minute j, let $x_j \in \mathbb{R}^3$ have coordinates corresponding to price of futures for S&P500, DJIA and NASDAQ.
- Fix a time t, and consider the dataset consisting of $x_{t-50}, \ldots, x_t$.
- Compute norm of persistence diagram. Consider the 'exponential moving average/variance' of this quantity.
- Let EMA_1 be the first norm, $\mathrm{EMVar}_1 = 0$.
- Define inductively

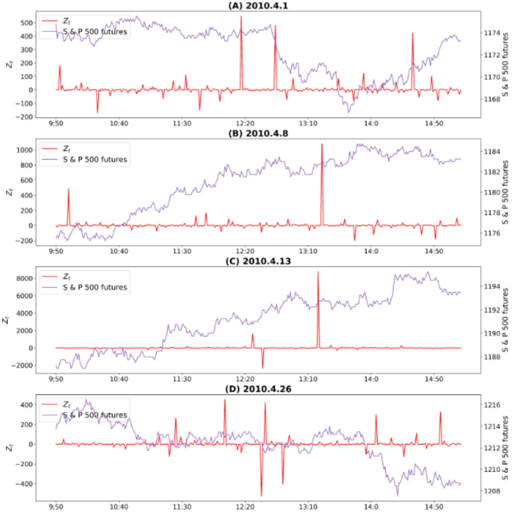
$$\begin{aligned} \delta_i &:= i\text{th Norm} - \mathrm{EMA}_{i-1}, \\ \mathrm{EMA}_i &:= \mathrm{EMA}_{i-1} + \alpha \cdot \delta_i, \\ \mathrm{EMVar}_i &:= (1-\alpha)(\mathrm{EMVar}_{i-1} + \alpha + \delta_i^2). \end{aligned}$$

- $Z_t = \frac{Y_t - \mathrm{EMA}_{t-1}}{\mathrm{EMVar}_{t-1}}$.

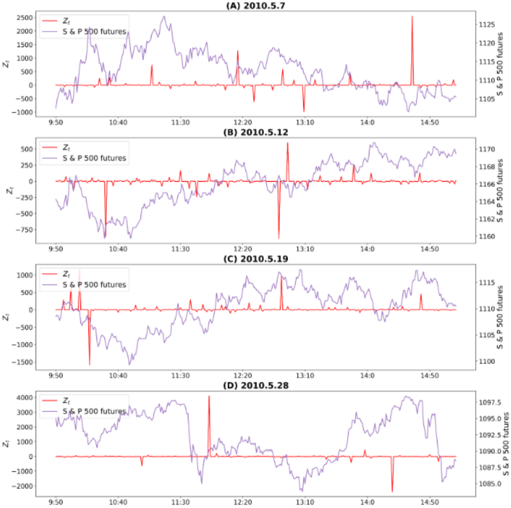
Surely not more graphs?!?



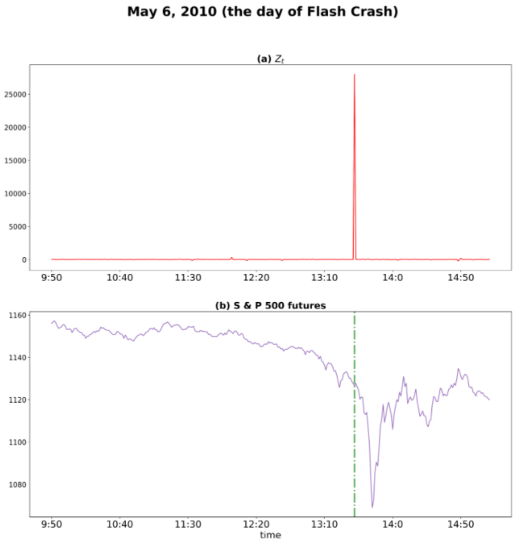
Pre-Event Graphs



Post-Event Graphs



I can't believe it's graphs



References I

- Marian Gidea, *Topology data analysis of critical transitions in financial networks*, 2017.
- Wonse Kim, Younng-Jin Kim, Gihyun Lee, and Woong Kook, *Investigation of flash crash via topological data analysis*.
- Miguel A. Ruiz-Ortiz, José Carlos Gómez-Larrañaga, and Jesús Rodríguez-Viorato, *A persistent-homology-based turbulence index and some applications of tda on financial markets*, 2023.