

Market crash and phase transitions

Y

2025

1 Introduction

We want to study the log returns of the S&P 500 index, with particular attention to what happened to the market crashes of 2000, 2008 and 2020 as we expect to witness a phase transition similar to what happens in physical systems.

2 Theoretical framework

A closing price and an opening price are associated to every index, where the closing price corresponds to the price of the index at the end of the day and similarly the opening price corresponds to the price of the index at the start of the day, so it is also natural to consider time intervals of a day. However, the closing price of a day does not correspond to the opening price of the following day, as when the market is not open the prices continue to change in relation to external effects.

3 Log returns

We can define the log returns $r(t)$ of the day t as the log difference between the closing price of that day and the one of the day before

$$r(t) = \log \text{Close}(t) - \log \text{Close}(t - 1).$$

The shaded areas in Fig. ?? correspond to the market crashes time interval. We can now plot Fig. ??

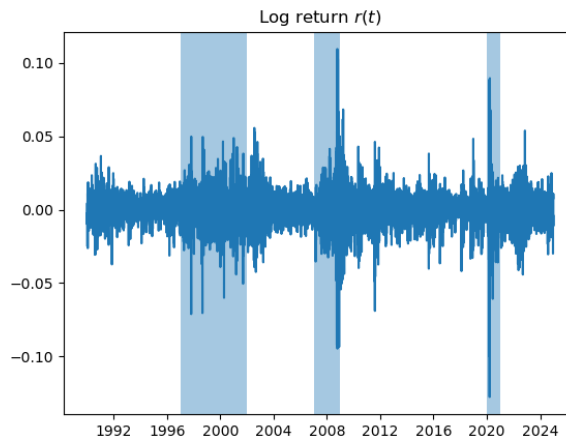


Figure 1: Log returns for the S&P500 index

the absolute value of the log returns to highlight the peaks reached. Lastly, we present the histogram of the log returns Fig. ?? which shows the probability of each return. We compare it with a gaussian pdf with the same mean and variance and we observe an higher peak and fatter tails of our data.

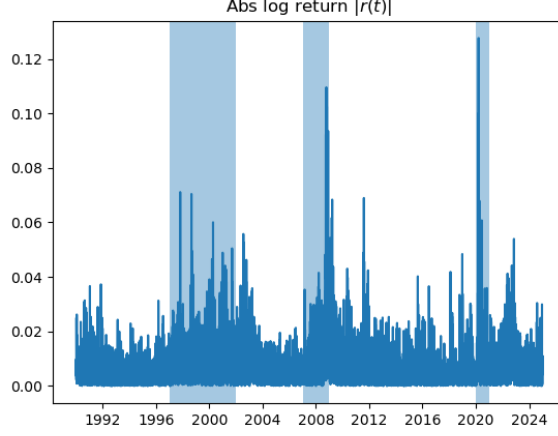


Figure 2: Absolute value of log returns for the S&P500 index

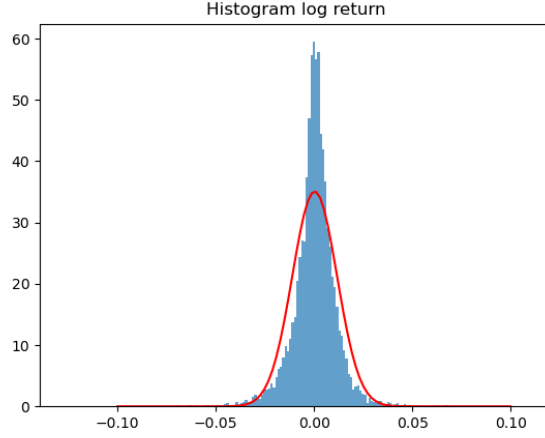


Figure 3: Histogram of the log returns of S&P500 index

4 Power law fit of the tails

We now want to study the tails of our distribution and quantitatively measure the deviation from the normal distribution; however, we cannot study directly our distribution, instead we will work around the problem. We start by ordering our events $\{x_k\}_{k=1,\dots,N}$, which correspond to the absolute value of the log returns, in decreasing order such that

$$x_1 \geq \dots \geq x_N.$$

The probability that x_k is greater than the following values is then

$$P_{>}(x_k) = \frac{k}{N}.$$

The idea justifying this result is the following: the probability is the ratio between favorable cases and all the cases (N). Because the list is ordered, the favorable cases correspond to the events with index greater than k and we have $N - k$ of them. The probability $P_{>}(x)$ is also called the Complementary Cumulative Distribution Function (CCDF).

Why do we need it? Suppose that our (unknown) distribution $P(x)$ has tails that follow a power law

$$P(x) \propto |x|^{-\alpha-1}.$$

The CDF then corresponds to

$$P_{>}(x) = P(y \geq x),$$

which can be evaluated by integrating our PDF $P(x)$. The integral of a power law is known so we get

$$P_{>}(x) \propto |x|^{-\alpha}.$$

We can then get the exponent α of our PDF by fitting the CCDF.

In our numerical analysis Fig. ?? we find that the exponent is

$$\alpha = 4.1.$$

The theory sets extreme values for this exponent: α cannot be lower than 2 as it would mean that our distribution has infinite variance, which is not physical; as α grows, then, the distribution resemble more a gaussian one $\alpha \rightarrow \infty$. In general one expect α to be upper limited by 4 as in this case variance is finite, but higher moments can still be infinite.

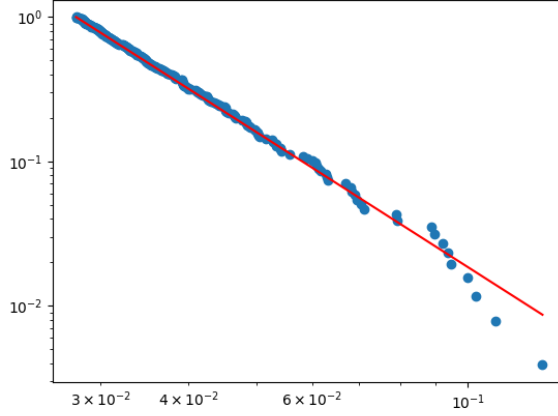


Figure 4: Fit of abs log returns of S&P500 index

5 Autocorrelations and volatility

Given our log return values $r(t)$ we can define the autocorrelations as

$$C(\tau) = \sum_t \langle r(t)r(t+\tau) \rangle.$$

We observe that these correlations decay very fast; for example, we report their behavior up to $\tau = 10$ days Fig. ??.

On the other hand, we can study the volatility, defined as the autocorrelations of the absolute log returns

$$C(\tau) = \sum_t \langle |r(t)||r(t+\tau)| \rangle.$$

As we can see from Fig. ??, volatility decays much slower compared to the autocorrelations, signaling that the market behavior is not totally random, but there is an underlying structure.

6 Approach to crash

We now evaluate the rolling tail exponent α for a 2 years window and construct a time series by moving this window of two months and re-evaluating the exponent. We then compare this series with the closing prices of the S&P 500 index Fig. ?. We expect that at a market crash the system behaves like at a critical point of a phase transition, thus becoming strongly correlated. We measure these correlations through α , so we expect the rolling tail exponent to decrease before the crash, because as α decreases the

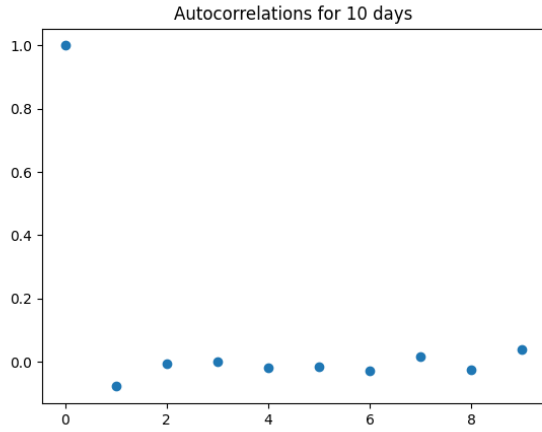


Figure 5: Autocorrelations of S&P500 index

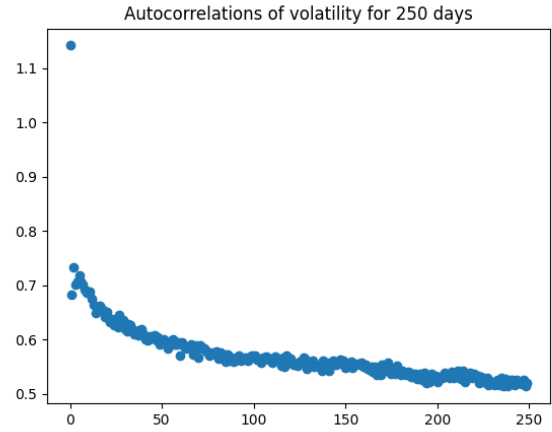


Figure 6: Volatility of S&P500 index

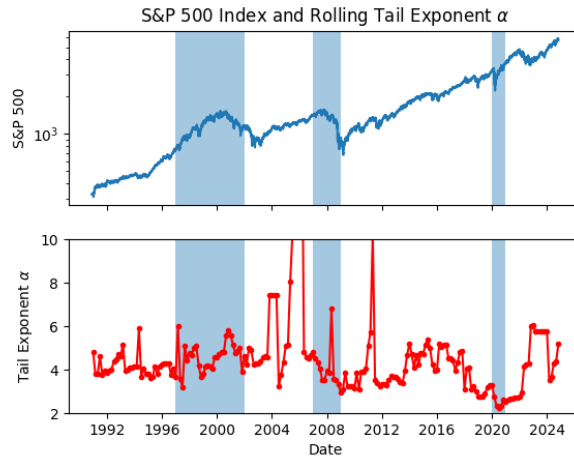


Figure 7: Rolling tail exponent compared with the S&P500 index

tails fatten and as the tails fatten extreme events become more likely: this is what α tells us in general, not only during periods of market crash. Let us discuss the role of α , the indicator of instability during the market crashes. Before the start of the dot-com bubble (1998), α rapidly decreases as the market is wildly unpredictable; α then rises during the crash (1999-2002). We see the same behavior during the crash of 2008: during the years preceeding the crash (2007), α starts to decrease and keeps doing it during 2008 and 2009. The COVID crash instead is different, as α does not decrease before the crash but only during it; this is due to a different kind of market crash. The first and second crashes were infact due to a bubble burst, while the last crash was due to an external event, a global pandemic.