Modeling Maxima with autoregressive conditional Fréchet model

Z. Zhao, Z. Zhang, R. Chen

December 10, 2020

Published on Journal of Econometrics (2018)

Presented by Liujun Chen

Table of Contents

Introduction

Maximum observations, as the representation of extreme behavior, are of particular interest.

- mutual fund managers: the potential maximum daily loss across all stocks in their managed portfolio
- high-frequency traders: the level of potential intra-day maximum loss

Introduction

Most applications of EVT focus on modeling extreme events in time series with a static approach under equilibrium distribution.

- The behavior of the underlying time series may change through time.
- Financial time series tends to exhibit structural changes and time-varying dynamics such as volatility clustering.

Introduction

Model time series of maxima $\{Q_t\}$, where $\{Q_t\}$ is a univariate time series of maxima based on a set of underlying financial time series $\{X_{it}\}_{i=1}^{p}$. There are mainly two types of Q_t.

- time series of cross-sectional maxima e.g. modeling the maximum daily loss across a group of stocks in a portfolio
- the time series of intra-period maxima e.g. intra-day maxima of high-frequency trading losses that occur on the same day.

Table of Contents

Dynamic Model

- A dynamic GEV framework: a conditional evolution scheme is designed for the parameters (μ_t, σ_t, ξ_t) of GEV, so that time dependency of $\{Q_t\}$ can be captured.
- Due to the heavy-tailedness of financial data, $\{Q_t\}$ can marginally can be well modeled by Fréchet distribution.
- $$Q_t = \mu_t + \sigma_t Y_t^{1/\alpha_t},$$

  where Y_t is a sequence of i.i.d. unit Fréchet random variables.

Should α_t be dynamic?

- Ad-hoc moving-window GEV analysis on the cross-sectional maxima of negative daily log-returns of S&P 100 index
- The observation period is from January 1, 2000 to December 31, 2014 with 3773 trading days.
- For each trading day t, we record the maximum daily loss across the 100 stocks and denote it by Q_t, Hence the time series $\{Q_t\}$ has 3773 observations
- For each $500 \leq t \leq 3273$, a GEV model is fitted using $\{Q_k\}_{k=t-499}^{t+500}$.

Moving Window Estimatio of Tail Index α

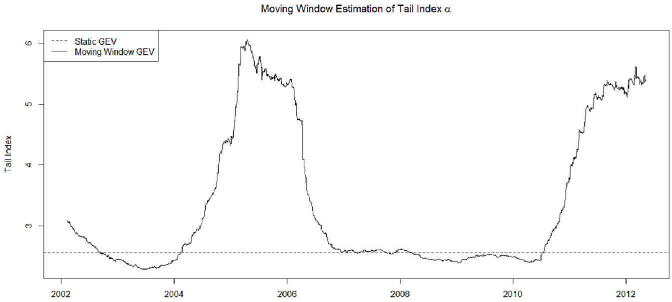


Fig. 1. Tail index $\hat{\alpha}_t$ estimated by moving window of size 1000 (solid curve) v.s. tail index $\hat{\alpha}$ estimated by the static GEV model based on total observations (dashed line).

Model Specification

Specifically, the autoregressive conditional Fréchet(AcF) model assumes the form

$$\log \sigma_t = \beta_0 + \beta_1 \log \sigma_{t-1} + \eta_1(Q_{t-1})$$
$$\log \alpha_t = \gamma_0 + \gamma_1 \log \sigma_{t-1} + \eta_2(Q_{t-1}),$$

where $\beta_1, \gamma_1 \geq 0$. The log transform is used to ensure the positivity of the parameters.

Assume that η_1 is a continuous increasing function and η_2 is a continuous decreasing function of Q_{t-1}. Choose η_1, η_2 to be the simple exponential function $a_0 \exp(-a_1 x)$.

Model Specifiction

For the rest of the paper, we consider the following model:

$$Q_t = \mu + \sigma_t Y_t^{1/\alpha_t},$$

$$\log \sigma_t = \beta_0 + \beta_1 \log \sigma_{t-1} - \beta_2 \exp(-\beta_3 Q_{t-1}), \qquad (*)$$

$$\log \alpha_t = \gamma_0 + \gamma_1 \log \sigma_{t-1} + \gamma_2 \exp(-\gamma_3 Q_{t-1}),$$

where $\{Y_t\}$ is a sequence of i.i.d. unit Fréchet random variables, $0 \leq \beta_1 \neq \gamma_1 < 1$, $\beta_2, \beta_3, \gamma_2, \gamma_3 > 0$.

Remark 1

- AcF can be easily extended to include q_1 autoregressive terms of $\log \sigma_t$ and $\log \alpha_t$, and q_2 lagged terms of $\eta(Q_t)$, similar to that of GARCH(q_1, q_2) model.
- Similar theoretical properties can be derived and similar estimation procedures can be used.
- Our empirical experience shows that the extension does not necessarily improve the performance of the model.

Stationarity and ergodicity

The model can be written as

$$\log \sigma_t = \beta_0 + \beta_1 \log \sigma_{t-1} - \beta_2 \exp(-\beta_3(\mu + \sigma_{t-1} Y_{t-1}^{1/\alpha_{t-1}})),$$

$$\log \alpha_t = \gamma_0 + \gamma_1 \log \sigma_{t-1} + \gamma_2 \exp(-\gamma_3(\mu + \sigma_{t-1} Y_{t-1}^{1/\alpha_{t-1}})),$$

Hence, $\{\sigma_t, \alpha_t\}$ form a homogeneous Markov chain in $\mathbb{R}^2$.

Theorem (1)

For an AcF with $\beta_2, \beta_3, \gamma_2, \gamma_3 > 0$, $\beta_0, \gamma_0, \mu \in \mathbb{R}$ and $0 \leq \beta_1 \neq \gamma_1 < 1$, the latent process $\{\sigma_t, \alpha_t\}$ is stationary and geometrically ergodic.

AcF under a factor model setting

We illustrate that the limiting form of maxima Q_t under a general factor model framework leads to an AcF model.

- Assume $\{X_{it}\}_{i=1}^{p}$ follow a general factor model,

$$X_{it} = f_i(Z_{1t}, \ldots, Z_{dt}) + \sigma_{it}\varepsilon_{it},$$

- $\{X_{it}\}_{i=1}^{p}$ are observed time series at time t, $\{Z_{1n}, \ldots, Z_{dn}\}$ consist of observed and unobserved factors, $\{\varepsilon_{it}\}_{i=1}^{p}$ are i.i.d. random noises that are independent with the factors $\{Z_{id}\}_{i=1}^{d}$.
- Assume that $\{\varepsilon_{it}\}_{i=1}^{p}$ are in the domain of attraction of the Fréchet distribution.

Other Assumptions for the factor model

- $\sup_{1\le p<\infty} \sup_{1\le i\le p} f_i(Z_{1t},\ldots,Z_{dt}) < \infty$ a.s.
- $\lim_{p\to\infty} \sum_{i=1}^{p} \sigma_{it}^{\alpha_t} = \infty$
- $\lim_{p\to\infty} \sup_{1\le i\le p} \frac{\sigma_{it}^{\alpha_t}}{\sum_{j=1}^{p} \sigma_{it}^{\alpha_t}} = 0$

Asymptotic conditional distribution of Q_t

Proposition (1)

Given F_{t-1}, denote $a_{pt} = 0$ and $b_{pt} = (\sum_{j=1}^{p} \sigma_{it}^{\alpha_t})^{1/\alpha_t}$, we have, as $p \to \infty$,

$$\frac{Q_t - a_{pt}}{b_{pt}} \xrightarrow{d} \exp(-x^{-\alpha_t}).$$

Table of Contents

MLE

- denote all the parameters in the model by $\theta = (\beta_0, \beta_1, \beta_2, \beta_3, \gamma_0, \gamma_1, \gamma_2, \gamma_3, \mu)$
- denote $\Theta_s = \{\theta | \beta_0, \gamma_0, \mu \in \mathbb{R}, 0 \leq \beta_1, \gamma_1 < 1, \beta_2, \beta_3, \gamma_2, \gamma_3 > 0\}$
- Assume that all allowable parameters are in Θ_s
- Denote the true parameter by $\beta_0 = (\beta_0^0, \beta_1^0, \beta_2^0, \beta_3^0, \gamma_0^0, \gamma_1^0, \gamma_2^0, \gamma_3^0, \mu_0)$

MLE

The conditional p.d.f. of Q_t given $(\mu_t, \sigma_t, \alpha_t)$ is

$$f_t(\theta) = f(Q_t|\sigma_t, \alpha_t) = \alpha_t \sigma_t^{\sigma_t} (Q_t - \mu)^{-(\alpha_t+1)} \exp\left\{-\sigma_t^{\alpha_t}(Q_t - \mu)^{-\alpha_t}\right\}.$$

Hence, by conditional independence, the log-likelihood function with observations $\{Q_t\}_{t=1}^n$ is

$$\begin{aligned} L_n(\theta) &= \frac{1}{n}\sum_{t=1}^{n} \log f_t(\theta) \\ &= \frac{1}{n}\sum_{i=1}^{n} \log \alpha_t + \sigma_t \log \sigma_t - (\alpha_t + 1)\log(Q_t - \mu) - \sigma_t^{\alpha_t}(Q_t - \mu)^{-\alpha_t} \end{aligned}$$

where $\{\sigma_t, \alpha_t\}$ can be obtained recursively through (*) with an initial value (σ_1, α_1).

Initial Value

- Notice here the true value of (σ_1, α_1), denoted as (σ_1^0, α_1^0), is an unknown since the state variables $\{\sigma_t, \alpha_t\}$ is a hidden process.
- Fortunately, with $0 \leq \beta_1, \gamma_1 < 1$, the influence of (σ_1, α_1) on future (σ_t, α_t) decays exponentially as t increases
- Hence its impact on parameter estimation will be minimum with a sufficiently large sample size.
- It will be shown that the consistency and asymptotic normality does not depend on the initial value.

Asymptotic Properties

Theorem (Consistency)

Assume the parameter space Θ is a compact set of Θ_s. Suppose the observations $\{Q_t\}_{t=1}^n$ are generated by a stationary and ergodic AcF with true parameter θ_0 and θ_0 is in the interior of Θ, then there exists a sequence θ_n of local maximizer of $L_n(\theta)$ such that $\theta_n \xrightarrow{p} \theta_0$ and $||\theta_n - \theta_0|| \leq \tau_n$ where $\tau_n = O_p(n^{-r}), 0 < r < 1/2$. Hence $\hat{\theta}_n$ is consistent.

Asymptotic Properties

Proposition (Asymptotic Uniqueness)

Denote $V_n = \{\theta \in \Theta | \mu \leq cQ_{n,1} + (1-c)\mu_0\}$ *where* $Q_{n,1} = \min_{1 \leq t \leq n} Q_t$, *under the conditions in the Theorem for consistency, for any fixed* $0 < c < 1$, *there exists a sequence of* $\hat{\theta}_n = \arg\max_{\theta \in V_n} \tilde{L}(\theta)$ *such that,* $\hat{\theta}_n \xrightarrow{P} \theta_0$, $||\theta_n - \theta_0|| \leq \tau_n$ *where* $\tau_n = O_p(n^{-r}), 0 < r < 1/2$, *and* $P(\hat{\theta}_n$ *is the unique global maximizer of* $\tilde{L}_n(\theta)$ *over* $V_n) \to 1$.

Asymptotic Properties

Theorem (Asymptotic Normality)

Under the conditions in Theorem for consistency, we have that $\sqrt{n}(\hat{\theta}_n - \theta_0) \xrightarrow{d} N(0, M_0^{-1})$, *where* $\hat{\theta}_n$ *is that in Theorem of consistency and* M_0 *is the Fisher Information matrix evaluated at* θ_0. *Further, the sample variance of plug-in estimated score function* $\left\{ \frac{\partial}{\partial \theta} \log f_t(\hat{\theta}_n) \right\}_{t=1}^{n}$ *is a consistent estimator of* M_0.

Table of Contents

Convergence of maxima in factor model

- we conduct numerical experiments to investigate the finite sample behavior of Q_t
- Specifically, we study the convergence of the marginal distribution of Q_t to its Fréchet limit under a one-time period factor model.
- To simplify notation, we drop the time index t in this section.

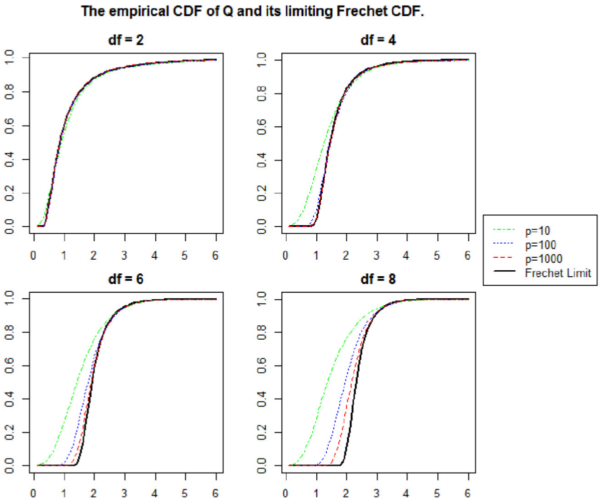
Convergence of maxima in factor model

We simulate data from the following one-factor linear model,

$$X_i = \beta_i Z + \sigma_i \varepsilon_i, \quad i = 1, \ldots, p,$$

- $Z \sim N(0,1)$ is the latent model,
- β_i are i.i.d. random coefficients generated from a uniform distribution $U(-2,2)$
- ε_i are i.i.d. t-distributions with degree of freedom v
- σ_i are i.i.d. random variables generated from a mixture of uniform distribution $\frac{1}{2}U(0.5,1.5) + \frac{1}{2}U(0.75,1.25)$

Finite sample empirical distribution of the maxima Q and its corresponding Fréchet limit



AcF estimation for conditional VaR of maxima

- temporal approximation ability of AcF to the maxima $\{Q_t\}$ process from a general factor model
- Model the $\{Q_t\}$ process using AcF and calculate the corresponding $cVaR_t^q$ for Q_t using the fitted AcF.
- $cVaR_t^q$ is defined as the $1-q$ extreme quantile of Q_t given all past information $\mathcal{F}_{t-1}$

Model Setting

- Simulate $\{Q_t\}$ process from a similar one-factor linear model:

$$X_i = 0.09(\beta_i Z_t + \sigma_i \varepsilon_{it}), \quad i = 1, \dots, p,$$

- $p = 100$

- $v_t = \gamma_0 + \gamma_1 \log v_{t-1} + \gamma_2 \exp(-\gamma_3 Q_{t-1})$ with $(\gamma_0, \gamma_1, \gamma_2, \gamma_3) = (-0.1, 0.9, 0.3, 5)$

Model Evaluation

- First fit AcF based on the training set $\{Q_t\}_{t=1}^{T_1}$.
- Then using fitted AcF, calculate $cVaR_t^q$ for each Q_t on the test set $\{Q_t\}_{t=T_1+1}^{T_1+T_2}$
- The true $\{Q_t\}_{t=T_1+1}^{T_1+T_2}$ are then compared with the $\left\{cVaR_t^q\right\}_{t=T_1+1}^{T_1+T_2}$ and the number of violations is recorded.
- A violation happens when the observed daily maxima Q_t is larger than the corresponding $cVaR_t^q$ given by AcF.
- If AcF approximates the tail behavior of $\{Q_t\}$ process well, the expected proportion of violations in the test set should be close to q.

Model Evaluation

- Also, assess the goodness of approximation by calculating the correlation between the true process v_t and the estimated process $\hat{\alpha}_t$ by AcF
- $T_1 = 1000, 2000, 5000$ $T_2 = 100$ and $q^0 = 0.1, 0.05, 0.01$
- repeat 500 times
- calculate the average violation times $\bar{q}$, mean and the median of the correlation.

Simulation Results

Table 1
The performance of AcF on approximation of 1-day conditional VaR for $\{Q_t\}$ process with independent errors ε_{it} and the correlation between the true tail index and the one estimated by AcF.

T_1	$\bar{q}\,(q^0 = 0.1)$	$\bar{q}\,(q^0 = 0.05)$	$\bar{q}\,(q^0 = 0.01)$	mean cor.	median cor.
1000	0.095	0.049	0.012	0.871	0.928
2000	0.096	0.049	0.012	0.909	0.952
5000	0.097	0.051	0.012	0.947	0.973

Performance of the MLE

To study the finite sample performance of the MLE, simulate the data from an AcF with the following parameters $(\beta_0, \beta_1, \beta_2, \beta_3, \gamma_0, \gamma_1, \gamma_2, \gamma_3, \mu) = (-0.050, 0.96, -0.051, 6.68, -0.068, 0.89, 0.33, 5.33, -0.069)$

Simulation Result

Table 4

Numerical results for performance of MLE with sample size 1000, 5000, 10 000. Mean and S.D. are the sample mean and standard deviation of the MLE's obtained from 500 simulations. 90% C.I. reports the coverage rate of the 90% C.I. constructed from the estimated Fisher Information matrix; 95% C.I. and 99% C.I. report corresponding coverage rates.

$N = 1000$	γ_0	γ_1	γ_2	γ_3	β_0	β_1	β_2	β_3	μ
True value	−0.068	0.890	0.330	5.33	−0.050	0.960	−0.051	6.68	−0.069
Mean	−0.060	0.884	0.346	6.28	−0.051	0.956	−0.054	5.88	−0.066
S.D.	0.029	0.028	0.058	1.93	0.028	0.019	0.023	3.25	0.011
90% *C.I.*	81	82	90	91	85	81	75	78	88
95% *C.I.*	84	88	93	94	87	87	79	80	95
99% *C.I.*	88	92	97	97	95	94	87	85	98
$N = 5000$	γ_0	γ_1	γ_2	γ_3	β_0	β_1	β_2	β_3	μ
True value	−0.068	0.890	0.330	5.33	−0.050	0.960	−0.051	6.68	−0.069
Mean	−0.066	0.889	0.332	5.52	−0.051	0.959	−0.052	6.53	−0.069
S.D.	0.014	0.012	0.029	0.88	0.012	0.008	0.009	1.83	0.005
90% *C.I.*	88	87	90	85	88	87	88	87	86
95% *C.I.*	92	96	93	94	92	91	93	93	94
99% *C.I.*	95	99	98	99	98	98	97	97	99
$N = 10\,000$	γ_0	γ_1	γ_2	γ_3	β_0	β_1	β_2	β_3	μ
True value	−0.068	0.890	0.330	5.33	−0.050	0.960	−0.051	6.68	−0.069
Mean	−0.067	0.890	0.330	5.44	−0.050	0.960	−0.051	6.55	−0.069
S.D.	0.010	0.007	0.018	0.61	0.007	0.005	0.006	1.37	0.003
90% *C.I.*	90	88	88	85	89	89	86	89	90
95% *C.I.*	93	94	94	94	92	94	93	94	98
99% *C.I.*	98	100	100	99	97	98	98	98	99

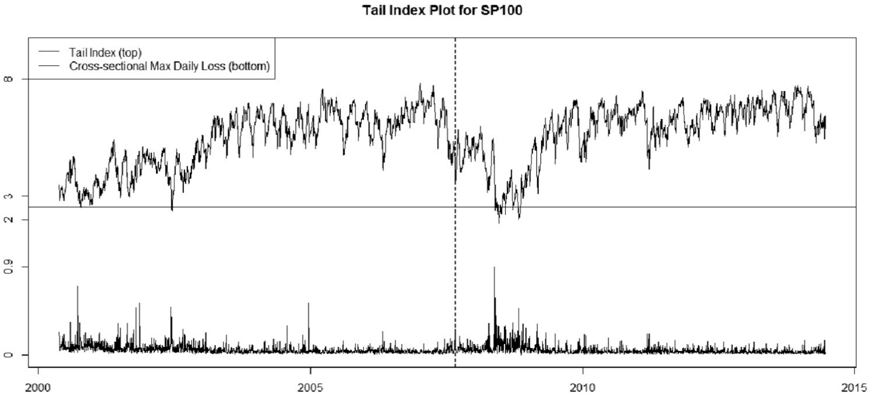
Table of Contents

Cross-sectional maxima of the negative daily log-returns of stocks

Table 5
MLE for cross-sectional maxima of negative daily log-returns for S&P100 (top) and DJI30 (bottom) from January 1, 2000 to December 31, 2014.

S&P100	γ_0	γ_1	γ_2	γ_3	β_0	β_1	β_2	β_3	μ
Mean	−0.068	0.890	0.328	5.33	−0.050	0.961	−0.051	6.68	−0.069
S.D.	0.014	0.013	0.063	1.27	0.006	0.004	0.0072	1.01	0.006
DJI30	γ_0	γ_1	γ_2	γ_3	β_0	β_1	β_2	β_3	μ
Mean	0.023	0.895	0.261	16.32	−0.052	0.964	−0.047	7.38	−0.059
S.D.	0.016	0.013	0.041	3.529	0.005	0.004	0.0066	0.813	0.006

Cross-sectional maxima of the negative daily log-returns of stocks



Intra-day maxima of 3-minute negative log-returns for USD/JPY foreign exchange rate

Table 6
MLE for intra-day maxima of 3-minute negative log-returns for USD/JPY from January 1, 2008 to June 26, 2013.

	γ_0	γ_1	γ_2	γ_3	β_0	β_1	β_2	β_3	μ
Mean	0.448	0.587	0.658	20.84	−0.120	0.890	−0.195	6.59	−0.051
S.D.	0.144	0.123	0.203	4.52	0.016	0.012	0.024	0.955	0.010

Intra-day maxima of 3-minute negative log-returns for USD/JPY foreign exchange rate

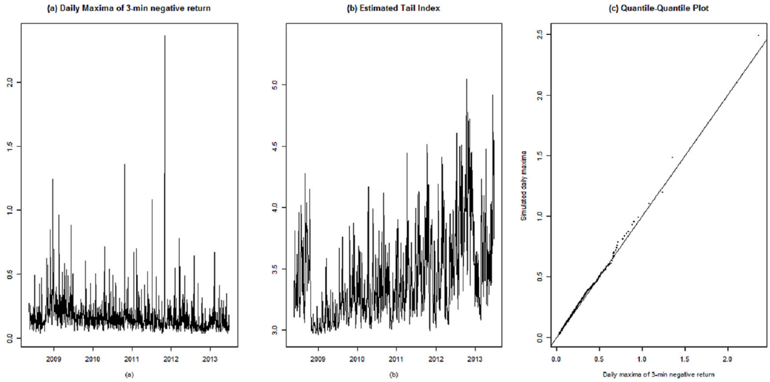


Fig. 7. (a) Daily maxima of 3-minute negative log-returns of USD/JPY from January 1, 2008 to June 26 2013; (b) Estimated tail index $\{\hat{\alpha}_t\}$ from the fitted AcF; (c) Quantile–quantile plot of real data and simulated data from the fitted AcF.