[P-ONE] Ongoing Roadmap

January 19, 2016

Contents

1 Introduction & Purpose

Tentative Title:
Asymptotic properties of extremum estimators and latent states in option price panels

Broad goals:

1. Derive asymptotic distribution (consistency/bias/asymptotic distribution) for the parameters of an AJD.
2. Find a better way to do specification testing.
3. Apply framework for a couple of well-established AJD models.

Specific goals:

1. Derive consistency for estimator with latent vol state in SV (Heston) model.
2. Extend set-up to cover:
    - 1 noisy option price;
    - noisy option price panels;
    - several latent states states/extensions of SV model;

Roadmap:
Tentative: Express $v_t(C, \theta)$ analytically $\longrightarrow$ compute derivative of moment condition analytically $\longrightarrow$ derive asymptotic distribution for a generic extremum criterion function.

2 Literature Review

2.1 Jarrow & Kwok (JoEctrcs, 2015) - Specification tests of calibrated option pricing models -

This paper proposes sets of assumptions under which the consistency and asymptotic distribution for estimators of option pricing model parameters can be established.[1] Most results are directly applicable for the Black-Scholes option pricing (uni-variate option pricing model with implied volatility as the estimated parameter).

Under the null assumption of a correctly specified model, statistical tests for misspecification are formulated for both the case in which the parameter estimate is obtained by inverting the option pricing function (e.g., backing out Black-Scholes implied volatility) and the case in which the parameter is estimated using a non-linear least squares procedure aimed at minimizing model pricing errors over a time-series of option prices.

Some of the assumptions needed for the asymptotics (notably in the set-up in which the option pricing function is inverted to obtain parameter estimates) are difficult to check, even in the univariate Black-Scholes set-up.

2.1.1 Set-up & Notation

We observe a panel of option prices over a cross-section of n options and a sample period $[0, T]$. Let $m_{it}(i = 1, \ldots, n; t = 1, \ldots, T)$ be the observed price of the ith option at time t, with strike price K_{it} and time to maturity τ_{it}. Suppose all options are written on the same underlying S_t, with dividend rate q_t. Let r_t denote the risk free interest rate. Collect all the observables in a vector $z_{it} = (K_{it}, \tau_{it}, S_t, q_t, r_t)$.

Modeler chooses a parametric option pricing model $M(\theta)$ indexed by a parameter θ so that $M_{it}(\theta) \coloneqq M(\theta; z_{it})$ is the theoretical model price for the option.

Observed option prices m_t are assumed to be noisy, in the following sense:

$$m_t = \mathring{m}_t + v_t. \tag{2.1}$$

v_t denotes the observation noise.

A necessary assumption for the results to be derived is:

<u>Assumption SM</u>: The option pricing function $M_t(\theta)$ is strictly monotone in θ.

[1] Sample counterparts of the asymptotic variance expressions would not constitute feasible estimators (?).

Error Minimization Calibration

θ is assumed constant across the entire sample[2].

The modeler defines a loss function, e.g., L_2 with equal weights:[3] $L_2 = \sum_{t=1}^{T}(M_t(\theta) - m_t)^2$, which it then minimizes w.r.t. θ to obtain an estimator based on the F.O.C.:

$$\begin{aligned} 0 &= \sum_{t=1}^{T} \underbrace{\frac{\partial M_t(\hat{\theta})}{\partial \theta}}_{\coloneqq \nabla_t(\hat{\theta})} (M_t(\hat{\theta}) - m_t) \\ &= \sum_{t=1}^{T} \nabla_t(\hat{\theta}) \left(\underbrace{M_t(\theta_0) + \nabla_t(\tilde{\theta})^\top (\hat{\theta} - \theta_0)}_{\text{Taylor expansion of } M_t(\hat{\theta})} - m_t \right) \\ &= \sum_{t=1}^{T} \nabla_t(\hat{\theta}) \left(M_t(\theta_0) + \nabla_t(\tilde{\theta})^\top (\hat{\theta} - \theta_0) - \underbrace{\mathring{m}_t + v_t}_{= m_t \text{ cf. (2.1)}} \right) \end{aligned} \tag{2.2}$$

For this derivation to hold it is required that:

Assumption D: The model $M_t(\theta)$ is continuously differentiable in θ so that $\nabla_t = \nabla_t(\theta) \coloneqq \frac{\partial M_t(\theta)}{\partial \theta}$ exists.

By further assuming:

Assumption H_0: No model misspecification, i.e. $\mathring{m}_t = M_t(\theta_0)$,

the F.O.C. in (2.2) further simplifies to:

$$\begin{aligned} 0 &= \sum_{t=1}^{T} \nabla_t(\hat{\theta}) \left(\nabla_t(\tilde{\theta})^\top (\hat{\theta} - \theta_0) - v_t \right) \\ \sum_{t=1}^{T} \nabla_t(\hat{\theta}) v_t &= \sum_{t=1}^{T} \nabla_t(\hat{\theta}) \nabla_t(\tilde{\theta})^\top (\hat{\theta} - \theta_0) \end{aligned} \tag{2.3}$$

Using the following additional assumptions:

Assumption SE: $\mathbb{E}(v_t|z) = 0$, $\forall t = 1, \ldots, T, \mathbb{P}$-a.s.
or Assumption WE: $\mathbb{E}(v_t|z_t) = 0$, $\forall t = 1, \ldots, T, \mathbb{P}$-a.s.

Assumption CV_1: $\mathrm{Var}(v_t|z_t) = \sigma^2$ and $\mathrm{Cov}(v_s, v_t|z_s, z_t) = 0$, $\forall t \neq s, \mathbb{P}$-a.s.
or Assumption CV_2: $\mathrm{Var}(v_t|z_t) = \sigma_t^2)$ and $\mathrm{Cov}(v_s, v_t|z_s, z_t) = 0$, $\forall t \neq s, \mathbb{P}$-a.s.

Assumption LC: The model $M_t(\theta)$ is strictly convex in θ in a neighborhood around θ_0.

[2]Latent state models are claimed to be accommodated within this framework (?).
[3]The cross-section $i = 1, \ldots, n$ dimension of the option panel is dropped in the derivations(?).

Assumption PL_a^{errmin}: A uniform weak LLN applies to $\hat{S}_{\nabla\nabla} := \frac{1}{T}\sum_{t=1}^{T}\nabla_t(\theta)\nabla_t(\theta)^\top$ (for $a=1$) or $\hat{S}_{\nabla\nabla} := \frac{1}{T}\sum_{t=1}^{T}\sigma_t^2\nabla_t(\theta)\nabla_t(\theta)^\top$ (for $a=2$), so that their probability limits $S_{\nabla\nabla}$ and $S_{\sigma^2\nabla\nabla}$ exist in a neighborhood Θ_0 of θ_0. Furthermore, $S_{\nabla\nabla}$ is invertible.

the following theorems can be stated:

Theorem 1: Under Assumptions SE, LC, and H_0, $\hat{\theta}$ is biased "in finite samples", i.e., $\mathbb{E}(\hat{\theta}) \neq \theta_0$.

Theorem 2: Under Assumptions WE and H_0, $\hat{\theta}$ is consistent in large samples, i.e., $\hat{\theta} \xrightarrow{p} \theta_0$ as $T \to \infty$.

Theorem 3: Under Assumptions WE, D, CV_a, PL_a^{errmin} (a=1,2) and H_0, $\hat{\theta}$ is asymptotically normally distributed:

$\sqrt{T}\left(\hat{\theta} - \theta_0\right) \xrightarrow{d} \mathrm{N}(0, V_a^{errmin})$ as $T \to \infty$, with asymptotic variance:

$$V_a^{errmin} = \begin{cases} \sigma^2 S_{\nabla\nabla}^{-1}, & \text{for } a = 1. \\ S_{\nabla\nabla}^{-1} S_{\sigma^2\nabla\nabla} S_{\nabla\nabla}^{-1}, & \text{for } a = 2. \end{cases}$$

Exact calibration

Exact calibration finds a solution for each t. The solution $\hat{\theta}_t$ varies across observations such that $\hat{\theta}_t$ satisfies for each t:

$$M_t(\hat{\theta}_t) = m_t = \mathring{m}_t + v_t \tag{2.4}$$

The estimator for θ_0 is then defined to be the sample mean of the calibrated parameters:

$$\begin{aligned} \bar{\theta} &= \sum_{t=1}^{T} \hat{\theta}_t && (2.5) \\ &= \sum_{t=1}^{T} M_t^{-1}(m_t) \\ &= \sum_{t=1}^{T} \left(\underbrace{M_t^{-1}(\mathring{m}_t) + \frac{\partial M_t^{-1}(\tilde{m}_t)}{\partial m_t}(m_t - \mathring{m}_t)}_{\text{Taylor expansion}} \right) \\ &= \sum_{t=1}^{T} \left(M_t^{-1}(\mathring{m}_t) + \frac{\partial M_t^{-1}(\tilde{m}_t)}{\partial m_t} v_t \right) && (2.6) \end{aligned}$$

Further assuming H_0 (no misspecification) leads to the simplified expression for the estimator:

$$\begin{aligned}\bar{\theta} &= \sum_{t=1}^{T} \theta_0 + \sum_{t=1}^{T} \frac{\partial M_t^{-1}(\tilde{m}_t)}{\partial m_t} v_t \\ &= \theta_0 + \sum_{t=1}^{T} \frac{\partial M_t^{-1}(\tilde{m}_t)}{\partial m_t} v_t \end{aligned} \tag{2.7}$$

The observation noise v_t is therefore transferred to the parameter space through this estimation approach.

3 Modeling