Some Traces of Discrete Mathematics in Mathematical Finance

Stefan Gerhold

TU Wien

Based on joint works with S. De Marco, P. Friz, P. Guasoni, J. Muhle-Karbe and W. Schachermayer

29 Jan 2013

Overview

- Introductory remarks
- **Topic 1:** Local volatility (saddle point asymptotics)
  De Marco, Friz, SG (RISK 2013)
- **Topic 2:** American option pricing (orthogonal polynomials)
  SG (Ann. Appl. Prob. 2011)
- **Topic 3:** Utility maximization under transaction costs (Lagrange inversion)
  SG, Guasoni, Muhle-Karbe, Schachermayer (Fin. Stoch., to appear)

Introductory remarks

- Price of a stock: Stochastic process S
- Discrete time $(S_t)_{t=0,\dots,T}$
- or continuous time $(S_t)_{0\leq t\leq T}$
- Values in $\mathbb{R}_{>0}$
- S_0 deterministic
- Fixed, given model; origin not to be explained here ($\rightarrow$ equilibrium theory)

Introductory remarks: Options

- **European call option:** Right to buy 1 share at time T for fixed price K
- K is called the strike (price)
- Payoff (value) at maturity T is

$$(S_T - K)^+ := \max\{S_T - K, 0\}$$

- American call option: May be exercised at any time $t \in [0, T]$, with payoff $(S_t - K)^+$.
- S is called the **underlying.**
- A call option is a basic **derivative** (security whose payoff depends on another security).

Introductory remarks: Portfolios

- Basic market model with two assets:
- Stock price process S_t
- Riskless bond $B_t = e^{rt}$, $r > 0$ constant
- Portfolio: Hold φ_t^0 units of bond, φ_t units of stock (predictable processes $\rightarrow$ don't look into the future)
- Value process: $V_t = \varphi_t^0 B_t + \varphi_t S_t$
- Values of φ_t^0, φ_t may be negative (borrowing / short selling)

Introductory remarks: Self-financing portfolios in discrete time

- Portfolio process $(\varphi_t^0, \varphi_t)_{t=1,\ldots,T}$
- (φ_t^0, φ_t) are holdings for period $[t-1, t[$
- Self-financing condition:

$$\varphi_t^0 B_t + \varphi_t S_t = \varphi_{t+1}^0 B_t + \varphi_{t+1} S_t$$

- At each t, rebalance portfolio without adding/removing funds
- Initial capital $V_0 := \varphi_1^0 B_0 + \varphi_1 S_0$ (price of setting up the strategy)

Introductory remarks: Option pricing

- Given: payoff profile H at time T
- Example: call option $H = (S_T - K)^+$.
- Goal: Replicate H, i.e., find portfolio process (φ^0, φ) with $V_T = H$.
- If possible: V_0 is price of H

Introductory remarks: Option pricing

- Original probability space $(\Omega, \mathcal{F}, P)$
- Compute prices as "risk-neutral" expectations

$$e^{-rT} E^*[H]$$

  with respect to a "risk-neutral" measure $P^* \approx P$
- $P^*[A]$, $A \in \mathcal{F}$, is not the real probability of A, but the price of the claim $H = \mathbf{1}_A$

Introductory remarks: Portfolio optimization

- Given: Initial capital V_0
- Goal: Generate "good" payoff
- Utility maximization:

$$E[U(V_T)] \to \max$$

- Utility function: increasing, concave, smooth, $U'(0) = \infty$, $U'(\infty) = 0$
- Examples: $U(x) = \log x$, $U(x) = \frac{x^{1-\gamma}}{1-\gamma}$, $1 \neq \gamma > 0$

Topic 1: Local volatility (option pricing)

- Idea: Construct a model that matches market call prices, then use it to price "exotic" options
- Given call price surface $C(K, T)_{K,T>0}$
- In practice: Interpolate from finitely many observed call prices
- Stock price model:

$$dS_t = \sigma_{loc}(S_t, t) S_t dW_t$$

- W standard Brownian motion; $W_{t+h} - W_t \sim \mathcal{N}(0, h)$
- $\sigma_{loc}(\cdot, \cdot)$ is a deterministic function such that

$$e^{-rT} E^*[(S_T - K)^+] = C(K, T) \quad \text{for all } K, T$$

Topic 1: Local volatility (option pricing)

- Consider the following setting:
- call surface $C(K, T)$ given by a *model*
- Determine the associated local vol model, i.e., determine $\sigma_{loc}(\cdot, \cdot)$
- Interesting: Asymptotics of $\sigma_{loc}(K, T)$ for large K
- Applications: Monte Carlo simulation, model risk

Topic 1: Local volatility (option pricing)

- Moment generating function of log-stock:

$$M(s,T) := E^*[S_T^s], \quad m := \log M$$

- Fourier representation of Dupire's formula ($k := \log K$):

$$\sigma_{\mathrm{loc}}^2(K,T) = \frac{2\partial_T C}{K^2 \partial_{KK} C} = \frac{2\int_{-i\infty}^{i\infty} \frac{\partial_T m(s,T)}{s(s-1)} e^{-ks} M(s,T)\,ds}{\int_{-i\infty}^{i\infty} e^{-ks} M(s,T)\,ds}$$

- Saddle point method: Leading terms are integrands evaluated at saddle point $\longrightarrow$ cancellation

General wing formula for local vol

- log moment generating function ($X_T = \log S_T$)

$$m(s,T) = \log E[\exp(s, X_T)]$$

- saddle point $\hat{s}(k,T)$

$$\left.\frac{\partial}{\partial s} m(s,T)\right|_{s=\hat{s}} = k$$

- "Lee type" wing formula for $k \to \infty$:

$$\boxed{\sigma^2_{\mathrm{loc}}(K,T) \approx \left.\frac{2\frac{\partial}{\partial T} m(s,T)}{s(s-1)}\right|_{s=\hat{s}(k,T)}}$$

- De Marco, Friz, SG (2012)

Heston model: Numerical example (left wing)

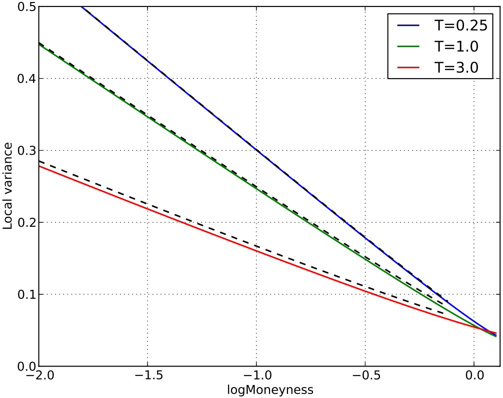


Abbildung: Local variance $\sigma^2_{\mathrm{loc}}(k, T)$ and our approximation in the Heston model.

Saddle point asymptotics

- Heston model, Kou model: mgf has singularity of the type

$$\exp\left(\frac{c}{s_+ - s}\right)$$

- s_+=critical moment (singularity of mgf)
- Heston: rigorous analysis (Friz, SG 2012).
  - saddle point $\approx s_+ - k^{-1/2}$
  - size of central contour part is $k^{-5/7}$
  - Tail estimate by ODE comparison (or: use explicit mgf)
- Merton jump diffusion: $s_+ = \infty$, mgf has double exponential blowup
- To do: give useful rigorous criteria for the validity of the saddle point approach

The variance gamma model

- mgf

$$M(s,T) = \left(\frac{1}{1-\theta\nu s - \frac{1}{2}\sigma^2\nu s^2}\right)^{T/\nu}$$

- saddle point

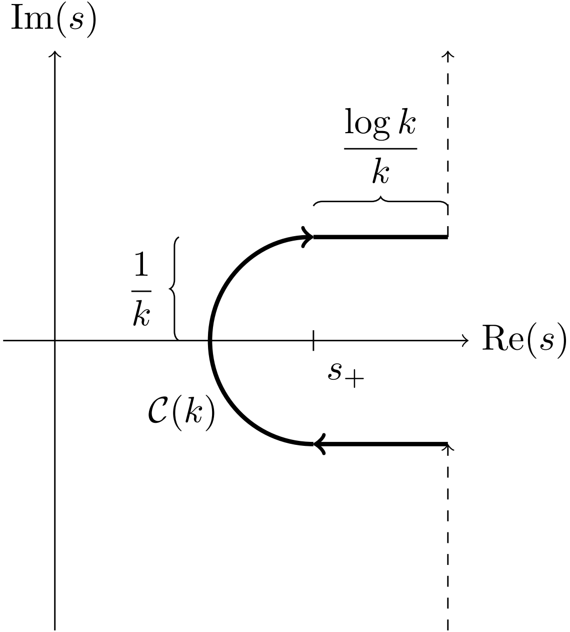
$$\hat{s} \approx s_+ - \frac{T}{\nu k}.$$

- Wing formula gives

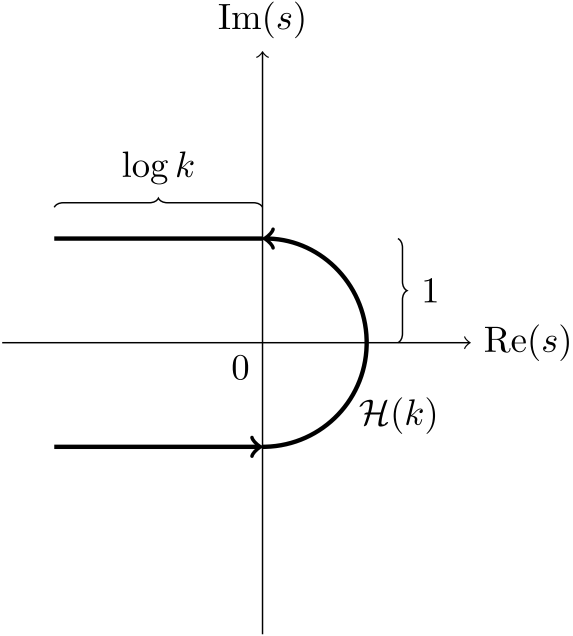
$$\sigma^2_{\mathrm{loc}}(k) \sim \frac{2\log(k/T)}{\nu s_+(s_+-1)}.$$

- Correct! Proof by Hankel contour integration
- Cf. Flajolet, Odlyzko (singularity analysis)
- Should work similarly for any regularly varying mgf

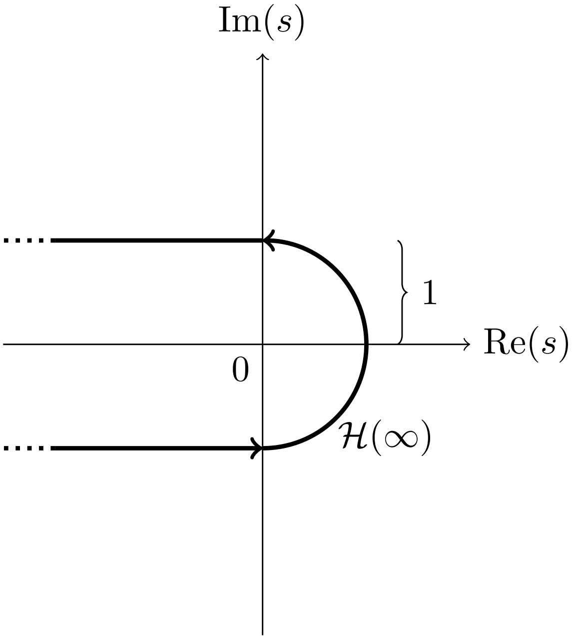
Hankel contours



Hankel contours



Hankel contours



Hankel contours

- Recall Fourier representation of local vol:

$$\sigma^2_{\mathrm{loc}}(K,T) = \frac{2\int_{-i\infty}^{i\infty} \frac{\partial_T m(s,T)}{s(s-1)} e^{-ks} M(s,T) ds}{\int_{-i\infty}^{i\infty} e^{-ks} M(s,T) ds}$$

- Denominator:

$$\frac{1}{2i\pi}\int_{-i\infty}^{i\infty} e^{-ks} M(s,T) ds \sim \frac{c_1 e^{-ks_+}}{\Gamma(c_2) k^{1-c_2}}.$$

- Numerator: Similar.
- $\Rightarrow$ wing formula $\sigma^2_{\mathrm{loc}}(K,T) \approx \left.\frac{2\frac{\partial}{\partial T} m(s,T)}{s(s-1)}\right|_{s=\hat{s}(k,T)}$ gives correct result.

Hankel contours

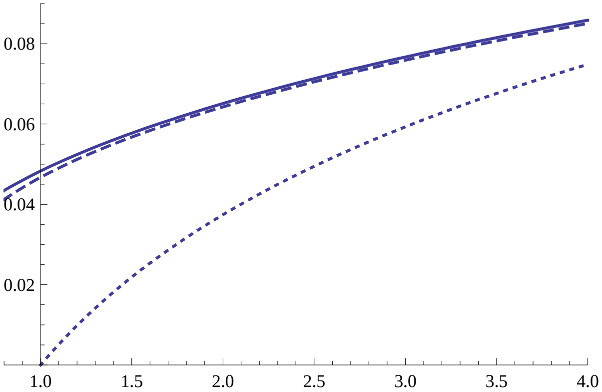


Abbildung: Local variance $\sigma^2_{loc}(k)$ for variance gamma model. **Puzzle:** Why is wing formula numerically superior??

Topic 2: American option pricing

- Markovian underlying S_t (multi-dimensional in general, but one-dimensional in our analysis)
- Exercise times $t_1, \ldots, t_m$
- discrete time steps
- Backward induction
- Successively price options with exercise times $t_n, \ldots, t_m$ $(n = m, \ldots, 1)$

Dynamic Programming

- $V_n(x) =$ value of the option with exercise times $t_n, \ldots, t_m$, at time t_n in state $S_{t_n} = x$.
- $h_n(x) =$ payoff function, if the option is exercised at time t_n.
- We have

$$V_n(x) = \max\{h_n(x), c_n(x)\},$$

  where $c_n(x)$ is the value of keeping the option.
- (Side remark: h_n may not have a closed form expression in practice.)

Dynamic Programming

- Continuation value

$$c_n(x) = E^*[V_{n+1}(S_{t_{n+1}}) \mid S_{t_n} = x]$$

- Backward induction

$$\begin{aligned} c_m(x) &= 0, \\ c_n(x) &= E^*[\max\{h_{n+1}(S_{t_{n+1}}), c_{n+1}(S_{t_{n+1}})\} \mid S_{t_n} = x]. \end{aligned}$$

- Option value at time t_0 is

$$\max\{h_0(S_{t_0}), c_0(S_{t_0})\}.$$

The Longstaff-Schwartz Algorithm

- Monte Carlo simulation of S_t
- Approximate continuation values

$$c_n(x) \approx \sum_{k=0}^{K} \beta_{nk} \psi_{nk}(x)$$

- Regression coefficients

$$\beta_n = (\beta_{n0}, \ldots, \beta_{nK})^{\mathrm{T}}$$

- Basis functions

$$\psi_n(x) = (\psi_{n0}(x), \ldots, \psi_{nK}(x))^{\mathrm{T}}$$

The Longstaff-Schwartz Algorithm

- n exercise dates, K basis functions, N Monte Carlo paths
- Approximation one: replace conditional expectations in the dynamic programming principle by projections on a finite set of functions taken from a suitable basis
- Approximation two: use Monte- Carlo simulations and least squares regression to compute the value function of the first approximation.

Convergence Results

- K basis functions, N Monte Carlo paths
- Partial results by Longstaff, Schwartz (2001) (2 exercise times)
- Clément, Lamberton, Protter (2002):
  - Almost sure convergence of first approximation, as K to infinity
  - K fixed: N to infinity, almost sure convergence of the MC procedure to the value function of approximation 1.
- Both $K, N \to \infty$: N must grow exponentially in K (Glasserman, Yu 2004).

Results of Glasserman and Yu (2004) and SG (2011)

- What is the highest K for which $MSE \to 0$ as $N, K \to \infty$?

- Geometric Brownian motion: $\sqrt{\log N}$ (Gl., Yu 2004)
- Brownian motion: $\log N$ (Gl., Yu 2004)

- Geometric Poisson process: $\log N / \log\log N$ (SG 2011)
- Geometric Gamma process: $\log N / \log\log N$ (SG 2011)
- Geometric Pascal process: $\log N / \log\log N$ (SG 2011)
- Geometric Meixner process: $\log N / \log\log N$ (SG 2011)

Results of Glasserman and Yu (2004) and SG (2011)

- Exponential increase in number of paths as number of basis functions increases
- Proofs depend on estimations of moments

$$E[\psi_{nj}(S_{t_n})\psi_{mk}(S_{t_m})], \qquad E[\psi_{nj}(S_{t_n})^2\psi_{mk}(S_{t_m})^2]$$

- ψ_{nk} is the k-th basis function at the n-th exercise opportunity.
- Simplifies greatly if
  - $(\psi_{nk}(S_{t_n}))_{0\leq n\leq m}$ is a martingale for each k
  - $(\psi_{nk})_{k\in\mathbb{N}}$ is orthogonal w.r.t. the distribution of S_{t_n}

Lévy-Sheffer Systems

- Sheffer Systems (Sheffer 1937, Meixner 1934): Given analytic functions f and u, what are the polynomials Q_k defined by

$$\sum_{k=0}^{\infty} Q_k(x)\frac{z^k}{k!} = f(z)\exp(xu(z))?$$

- Lévy-Sheffer Systems (Schoutens 2000): Define $Q_m(x,t)$ by

$$\sum_{k=0}^{\infty} Q_k(x,t)\frac{z^k}{k!} = f(z)^t \exp(xu(z)).$$

Lévy-Sheffer Systems

- Lévy-Sheffer Systems (Schoutens 2000): Define $Q_m(x,t)$ by

$$\sum_{k=0}^{\infty} Q_k(x,t)\frac{z^k}{k!} = f(z)^t \exp(xu(z)).$$

- Assume: f, u analytic at zero
- Assume: $1/f(u^{-1}(i\theta))$ is the characteristic function of an infinitely divisible distribution, defines a Lévy process X_t
- Defines $Q_m(x,t)$, polynomial in x
- Martingale property

$$E[Q_k(X_t,t) \mid X_s] = Q_k(X_s,s)$$

Examples of Lévy-Sheffer Systems

- Hermite polynomials, Brownian Motion
- Charlier polynomials, Poisson process
- Laguerre polynomials, Gamma process
- Meixner polynomials, Pascal process
- Meixner-Pollaczek polynomials, Meixner process

Martingale Properties

- Charlier, Laguerre, Meixner, Meixner-Pollaczek polynomials

$$
\begin{aligned}
E^*[C_k(N_t, t) \mid N_s] &= \left(\frac{s}{t}\right)^k C_k(N_s, s), \\
E^*[L_k^{(t-1)}(G_t) \mid G_s] &= L_k^{(s-1)}(G_s), \\
E^*[M_k(P_t; t, q) \mid P_s] &= \frac{(s)_k}{(t)_k} M_k(P_s; s, q), \\
E^*[P_k(H_t; t, \zeta) \mid H_s] &= P_k(H_s; s, \zeta),
\end{aligned}
$$

- Connect moments at different exercise times, useful in the analysis

Convergence Results (SG 2011)

- N paths, K basis functions
- S_t geometric Poisson process, $\psi_{nk}(x) = t_n^k C_k(\log x, t_n)$ (Charlier polynomials)
- Put $(u, v) = (10, 4)$.
- **If $N \geq K^{(u+\varepsilon)K}$, then the mean square error tends to zero.**
- **If $N \leq K^{(v-\varepsilon)K}$, then the mean square error tends to infinity.**
- For the geometric Gamma, Pascal, and Meixner process, replace (u, v) by $(8, 8)$, $(11, 7)$, and $(8, 8)$, respectively.

Proof Ingredients of the Convergence Results

- A formula of Zeng (1992) for linearization coefficients of Meixner-Pollaczek polynomials
- Linearization coefficients are the coefficients a_j in

$$\psi_{nk}^2 = \sum_{j=0}^{k} a_j \psi_{nj}.$$

- A classical formula connecting Laguerre polynomials with different parameters
- Estimate norm of the inverse of a tridiagonal matrix
- Estimate some involved sums

Topic 3: Utility maximization under transaction costs

Merton (1971, 1973): Consider a safe and a risky asset, following

$$dB_t/B_t = rdt, \quad dS_t/S_t = (\mu + r)dt + \sigma dW_t, \quad \mu, \sigma > 0$$

For an investor with constant relative risk aversion γ and arbitrary time horizon:

- Optimal policy: constant risky fraction $\pi_* = \mu/\gamma\sigma^2$, only depends on (r, μ, σ) through **mean-variance ratio** μ/σ^2
- Utility grows at certainty equivalent rate $\beta = r + \mu^2/2\gamma\sigma^2$
- Def. of risky fraction:

$$\pi_t = \frac{\varphi_t S_t}{\varphi_t^0 B_t + \varphi_t S_t}$$

Topic 3: Utility maximization under transaction costs

- Black-Scholes model for S^0, S as before
- But now, buy at **ask price** S, sell at **bid price** $(1-\epsilon)S$, where $\epsilon \in (0,1)$ is the width of the bid-ask spread.

Simplest possible benchmark optimization problem?

- Investment depends on time horizon $\Rightarrow$ infinite horizon
- Utility from consumption (Magill and Constantinidis (1976), Davis and Norman (1990), Shreve and Soner (1994)): Keep risky fraction in **no-trade region** $[\pi_-, \pi_+]$ around π_*
- π_-, π_+ characterized by a free boundary problem
- Janeček and Shreve (2004): First-order asymptotics via viscosity approach

Maximize the **certainty equivalent rate**

$$\liminf_{T\to\infty} \frac{1}{T}\frac{1}{1-\gamma} \log E\left[\left(\varphi^0_T + \varphi^+_T(1-\epsilon)S_T - \varphi^-_T S_T\right)^{1-\gamma}\right] \to \max!$$

With transaction costs:

- Simplest case: log-utility (γ "=" 1). Solution in Taksar et al. (1989), asymptotics in SG, Muhle-Karbe & Schachermayer (2011)
- No trade region $[\pi_-, \pi_+]$, optimal growth rate known via the solution of a one-dimensional equation
- $\gamma \neq 1$: Similar heuristic results of Dumas and Luciano (1991)
- Starting point of our analysis

Main results and Implications

Definition of the key extra parameter, the gap

Lemma

For small $\epsilon > 0$, there is a unique $\lambda > 0$ for which the solution to

$$w'(x) + (1-\gamma)w(x)^2 + \left(\frac{2\mu}{\sigma^2} - 1\right) w(x) - \gamma \left(\frac{\mu-\lambda}{\gamma\sigma^2}\right)\left(\frac{\mu+\lambda}{\gamma\sigma^2}\right) = 0$$

$$w(0) = \frac{\mu-\lambda}{\gamma\sigma^2},$$

also satisfies the following terminal condition:

$$w(\log(u_\lambda/l_\lambda)) = \frac{\mu+\lambda}{\gamma\sigma^2} \quad \textit{where} \quad u_\lambda/l_\lambda = \frac{1}{(1-\epsilon)}\frac{(\mu+\lambda)(\mu-\lambda-\gamma\sigma^2)}{(\mu-\lambda)(\mu+\lambda-\gamma\sigma^2)}$$

- Scalar equation for λ, implicit function theorem yields:

$$\lambda = \gamma\sigma^2 \left(\frac{3}{4\gamma}\pi_*^2 (1-\pi_*)^2\right)^{1/3} \epsilon^{1/3} + O(\epsilon).$$

Main results and Implications

Optimal policy

Theorem

It is a long-run optimal policy to keep the risky weight within the buy and sell bounds

$$\pi_{\pm} = \frac{\mu \pm \lambda}{\gamma \sigma^2} = \pi_* \pm \left(\frac{3}{4\gamma} \pi_*^2 (1 - \pi_*)^2 \right)^{1/3} \epsilon^{1/3} + O(\epsilon).$$

- By evaluating π_-, π_+ in terms of the trading prices, the no-trade region becomes symmetric around $\pi_* = \mu / \gamma \sigma^2$
- Boundaries known explicitly in terms of μ, σ, γ and the gap λ
- At the first order, same result as in Janeček and Shreve (2004), i.e., investment and consumption separate
- As without transaction costs, policy only depends on investment opportunities μ, σ via mean-variance ratio μ / σ^2

Main results and Implications

Trading volume

Theorem

***Relative turnover**, i.e., the number $d\|\varphi\|_t$ of shares traded as a fraction of the number $|\varphi_t|$ of shares held, has long-term average*

$$\lim_{T\to\infty} \frac{1}{T}\int_0^T \frac{d\|\varphi\|_t}{|\varphi_t|} = \left(1-\frac{\mu-\lambda}{\gamma\sigma^2}\right)\frac{\sigma^2}{2}\left(\frac{\frac{2\mu}{\sigma^2}-1}{(u_\lambda/l_\lambda)^{\frac{2\mu}{\sigma^2}-1}-1}\right) + \left(1-\frac{\mu+\lambda}{\gamma\sigma^2}\right)\frac{\sigma^2}{2}\left(\frac{1-\frac{2\mu}{\sigma^2}}{(u_\lambda/l_\lambda)^{1-\frac{2\mu}{\sigma^2}}-1}\right)$$

- Proposed in the empirical literature to measure trading activity
- Also determined completely in terms of μ, σ, γ and the gap λ
- In our model, real-world trading volume is a puzzle for the opposite reason as the equity premium: The implied risk aversion is too low

Lagrange inversion

- Goal: Asymptotics of $\lambda = \lambda(\varepsilon)$ as $\varepsilon \to 0$ ($\Rightarrow$ all other asymptotics)
- $w(\lambda, x)$ solution of

$$w'(x) + (1-\gamma)w(x)^2 + \left(\tfrac{2\mu}{\sigma^2} - 1\right) w(x) - \gamma \left(\tfrac{\mu-\lambda}{\gamma\sigma^2}\right)\left(\tfrac{\mu+\lambda}{\gamma\sigma^2}\right) = 0$$

$$w(0) = \tfrac{\mu-\lambda}{\gamma\sigma^2},$$

- Series expansion (W_{ij} simple functions of model parameters):

$$w(\lambda, x) = \frac{\mu - \lambda}{\gamma\sigma^2} + \sum_{i=1}^{\infty}\sum_{j=0}^{\infty} W_{ij} x^i \lambda^j,$$

- λ is determined by

$$w(\lambda, \log u(\lambda)/l(\lambda)) = (\mu + \lambda)/\gamma\sigma^2$$

Lagrange inversion

- Power series of w yields equations

$$\lambda^3 \sum_{i \geq 0} A_i \lambda^i = \varepsilon \sum_{i,j \geq 0} B_{ij} \varepsilon^i \lambda^j$$

- First coefficients:

$$A_0 = \frac{4}{3\mu\sigma^2(\gamma\sigma^2 - \mu)} \qquad \text{and} \qquad B_{00} = \frac{\frac{1}{2} 2\mu(\gamma\sigma^2 - \mu)}{\gamma^2\sigma^4}$$

- Divide by $\sum_{i \geq 0} A_i \lambda^i$, and take the third root $\Rightarrow$

$$\lambda = \varepsilon^{1/3} \sum_{i,j \geq 0} C_{ij} \varepsilon^j \lambda^j = \varepsilon^{1/3} \sum_{i,j \geq 0} C_{ij} (\varepsilon^{1/3})^{3i} \lambda^j \, .$$

- Can be solved for λ, as a power series in $\varepsilon^{1/3}$