Forecasting implied volatility in a risk management problem

Hirbod Assa[1] Mostafa Pouralizadeh[2] (Speaker)
Abdolrahim Badamchizadeh[3]

[1]University of Liverpool, Liverpool, England

[2]Allameh Tabataba'i University, Tehran, Iran

[3]Allameh Tabataba'i University, Tehran, Iran

Approaches & Goals

- Providing a risk management problem
- Modeling implied volatility with a machine learning approach to help improve the precision of contract pricing
- Removing the drawbacks of static arbitrage
- Taking care of the trade off between high bias and high variance during the fitting procedure

The importance of the subject

- We face with a pricing formula for an European call option in our risk management problem
- To price call options the volatilty should be correctly specified
- Especially after the crash of 1987, the volatilty has been no longer fixed and it varies directly by time to maturity and strike price.
- So, having an adequate model to describe the behavior of implied volatility is an area of concern for the recent decades

Previous studies

- M. Malliaris, L. Salchenberger (1996), Using neural networks to forecast the SP 100 implied volatility.
- Cont, Rama, and Jos Da Fonseca (2002), Dynamics of implied volatility surfaces.
- A. Alentorn (2004), Modelling the implied volatility surface.
- J. Gatheral and A. Jacquier (2014), Arbitrage-free SVI volatility surfaces.

Risk management problem

Risk of loss

- $(S_t)_{0<t<T}$:The value of a risky asset over the time interval $[0,\ T]$
- r: The rate of risk-free asset
- T: Expiration time

$$L = \left(S_0 - e^{-rT} S_T\right)^+$$

Agent's global position

- X: Contract
- $\pi\,(X)$:The price of contract

$$P = L - X + \pi\,(X) \quad , \quad 0 \leq X \leq L$$

Comonotonicity

An R^2-valued randome vector $X = (X_1,\ X_2)$ is called comonotonic if

$$P\left(X_1 \leq x_1,\ X_2 \leq x_2\right) = \min_{i \in \{1,2\}} P\left(X_i \leq x_i\right) \quad , \quad (x_1,\ x_2) \in R^2$$

In other words, X and Y are almost surly increasing function of $X+Y$

To avoid moral hazard

X and L-X are assumed to be comonotonic

- X an L-X are increasing functions of L
- The more loss we face, the more gain we have from the contract
- The more loss we face, the more distance there is between loss and gain

Valuation of the contract

Fundamnetal theorem of asset pricing

- No-arnitrage condition guarantees the existence of at least one equivalent martingale measures

Price of the contract

$$\pi(X) = \sup_{Q \in \Delta} E^{Q}(X)$$

- Q : An equivalent martingale measure
- Δ : A subset of all equivalent martingale measures

Value at Rist (VaR)

A measure of risk of an investment

$$VaR_{\alpha}(Y) = \inf \{ y \in R \,|\, F_Y(y) \geq \alpha \}$$

- Y : The risk of loss
- $F_Y(y)$: Cumulative distribution function of the risk
- α : Risk aversion of the agent

Global risk

$$\min_{\substack{0 \leq X \leq L \\ X,\ L-X \text{ are} \\ \text{comonotone}}} VaR_{\alpha}(L - X) + \sup_{Q \in \Delta} E^{Q}(X)$$

Solution

Theorem

An optimal solution X to the risk management problem is given by

$$X = \min\left\{ \left(S_0 - e^{-rT} S_T\right)^+ + \left(S_0 - e^{-rT} VaR_{(1-\alpha)}\left(S_T\right)\right)^+ \right\}$$

The form of the solution is given as $X = f\left(S_T\right)$, where

$$f\left(x\right) = \left(S_0 - e^{-rT} VaR_{1-\alpha}\left(S_T\right)\right)^+ + e^{-rT}\left(x - S_0 e^{-rT}\right)^+ \\ - e^{-rT}\left(x - VaR_{1-\alpha}\left(S_T\right)\right)^+$$

Strike prices

The following terms play the role of stike price in option pricing

- $K_1 = S_0 e^{-rT}$
- $K_2 = VaR_{1-\alpha}\left(S_T\right)$

Strategy

Main issues

1. The call options are correctly priced
2. Volatility should be carefully specified

- we consider a Black-Scholes strategy
- A polynomial learning algorithm is used to parametrize implied volatilty

Static arbitrage

Definition

A volatility surface is free from static arbitrage if and only if the following conditions are satisfied:

1. The surface is free from calendar spread arbitrage;
2. For a fixed time to matiurity, the volatility slice is free from butterfly arbitrage.

Parameterization of implied volatility

Stochastic Volatility Inspired

$$W^{SVI}_{imp}(x) = a + b(\rho(x-m) + \sqrt{(x-m)^2 + \sigma^2})$$

$$a \in \mathbb{R} \;,\quad b \geq 0 \;,\quad |\rho| < 1 \;,\quad m \in \mathbb{R} \;,\quad \sigma > 0$$

$$W^{SVI}_{imp} = \tau \sigma^2_{imp} \qquad , \qquad x = \log \frac{K}{F_{[t,\; t+\tau]}} = \log \frac{K}{e^{r\tau} S_0}$$

- W^{SVI}_{imp} :Total implied variance
- x : Moneyness

Model selection

Polynomial approach

$$w_\theta^{Q^2}(x) = \theta_0 + \theta_1 x + \theta_2 x^2$$
$$w_\theta^{Q^3}(x) = \theta_0 + \theta_1 x + \theta_2 x^2 + \theta_3 x^3$$
$$w_\theta^{Q^4}(x) = \theta_0 + \theta_1 x + \theta_2 x^2 + \theta_3 x^3 + \theta_4 x^4$$

Intuition

- Linear model results in under-fitting
- Add some polynomial features
- Quadratic model is adequate mathematically
- Implied volatility do not always resemble a quadratic form
- In case of low VIX (volatility index), we need higher degree polynomial
- High degree polynomial may cause the problem of over-fitting

Cost function

$$\hat{\theta} = \arg\min_{\theta} \frac{1}{m}\left(\sum_{i=1}^{m}\left(w_{\theta}^{Q^n}(x^{(i)}) - w^{(i)}\right)^2 + \lambda\sum_{j=1}^{n}\theta_j^2\right)$$

- m: Number of training example
- $w^{(i)}$:The vector of observed total implied variance in market
- λ: Regularization parameter

Intuition

- The penalty term keeps the parameters small to avoid over-fitting
- The regularization parameter control the trade-off between bias and variance
- It is more probable to face butterfly arbitrage in absence of the penalty term

Devision of training set

To improve fitting robustness, the training set data is randomly divided into three portions

1. The training set (60%)
2. The cross validation set (20%)
3. The test set (20%)

Algorithm

A simple pseudo code

1. Start with the Total implied variance and its corresponding moneyness for the OTM data
2. For a fixed value of λ estimate parameters using training set data
3. Compute training error and cross-validation error
4. Plot learning curve
5. If there is no effect of over-fitting and under-fitting, compute test error for each model

   Otherwise, plot validation curve and choose the optimum value of λ, then move back to step 2
6. Choose the model with the lowest test error

Table: Machine learning approach for S&P 500, traded on DEC 15, 2014.

Time to maturity	Expiry date	n	λ	The best pair
0.0136	12/20/2014	2 3 4	0.001 0.003 0.005	$n = 2$ $\lambda = 0.001$
0.0438	12/31/2014	2 3 4	0.3 0.001 0.003	$n = 3$ $\lambda = 0.001$
0.0684	01/09/2015	2 3 4	3 9.95 10	$n = 2$ $\lambda = 3$
0.0904	01/17/2015	2 3 4	0.01 0.05 0.1	$n = 2$ $\lambda = 0.01$
0.126	01/30/2015	2 3 4	0.05 3 3	$n = 2$ $\lambda = 0.05$
0.178	02/20/2015	2 3 4	3 3 10	$n = 3$ $\lambda = 3$

Durrleman 's function

- there is no butterfly arbitrage for ML Approach
- there is no butterfly arbitrage for SVI Model

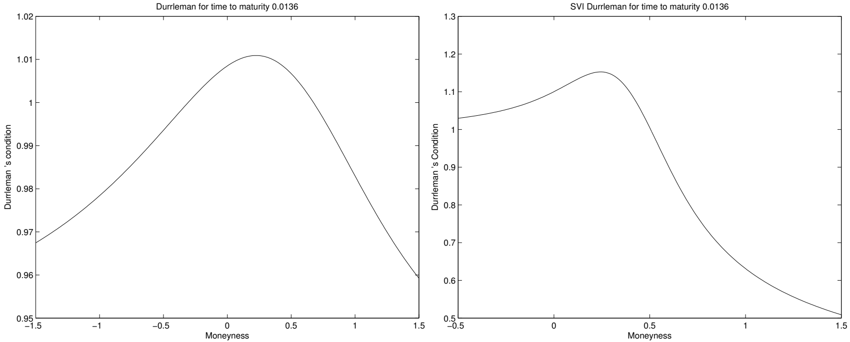


Figure: Durrleman function of ML and SVI, implemented for $\tau = 0.0136$

Durrleman 's function

- there is no butterfly arbitrage for ML Approach
- there is butterfly arbitrage for SVI Model

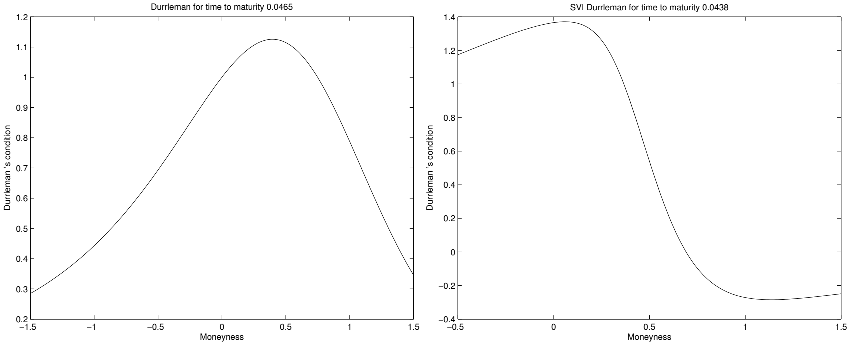


Figure: Durrleman funct4on of ML and SVI, implemented for $\tau = 0.0438$

Durrleman 's function

- there is no butterfly arbitrage for ML Approach
- there is butterfly arbitrage for SVI Model

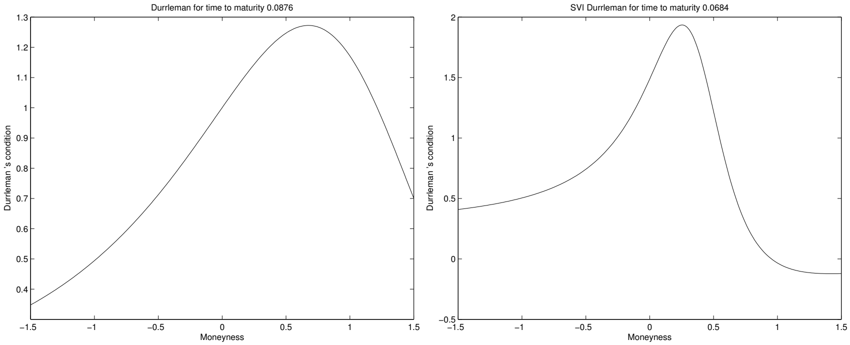


Figure: Durrleman function of ML and SVI, implemented for $\tau = 0.0684$

Durrleman 's function

- there is no butterfly arbitrage for ML Approach
- there is butterfly arbitrage for SVI Model

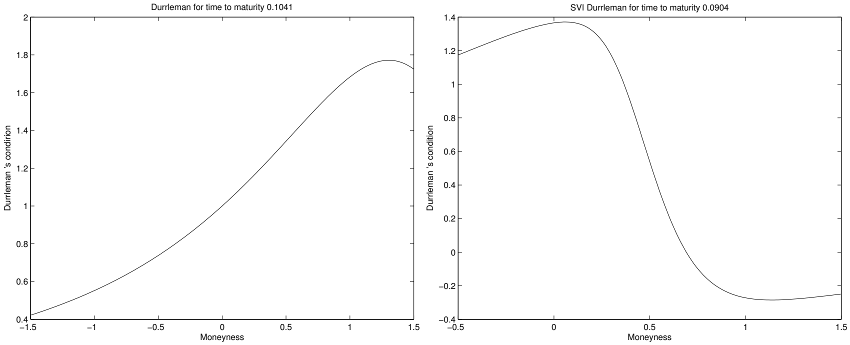


Figure: Durrleman function of ML and SVI, implemented for $\tau = 0.0904$

Durrleman 's function

- there is no butterfly arbitrage for ML Approach
- there is no butterfly arbitrage for SVI Model

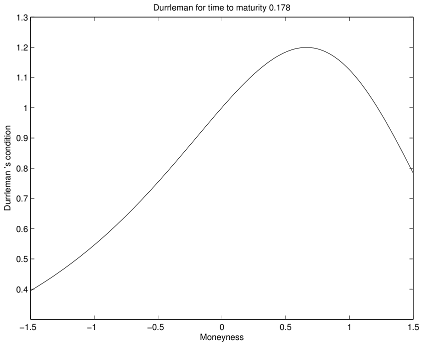


Figure: Durrleman function of ML and SVI, implemented for $\tau = 0.126$

Durrleman 's function

- there is no butterfly arbitrage for ML Approach
- there is butterfly arbitrage for SVI Model

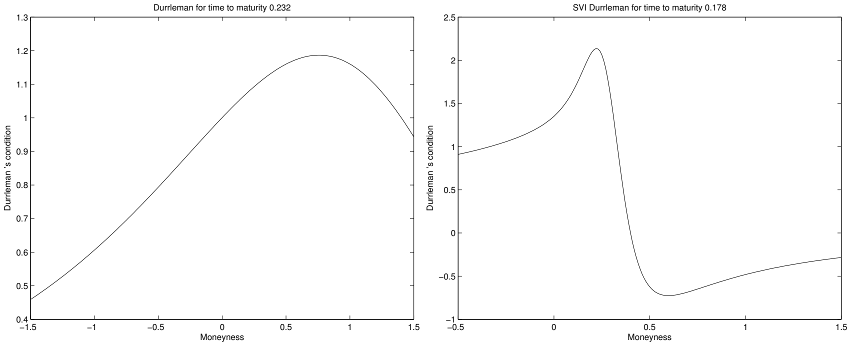


Figure: Durrleman function of ML and SVI, implemented for $\tau = 0.178$

Calendar spread plot

- there is no callendar spread arbitrage arbitrage for ML Approach

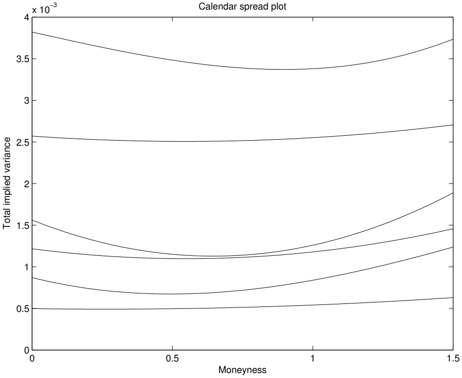


Figure: Total implied variance

The quadratic model

$$w_\theta^{Q^2}(x) = \theta_0 + \theta_1 x + \theta_2 x^2$$

Butterfly arbitrage

The quadratic model for less than one year time to maturity ($\tau < 1$) is free from butterfly arbitrage if

1. $\theta_1^2 - 4\theta_0\theta_2 + \theta_2 < 0$
2. $\frac{1}{4} < \theta_0 < 1$

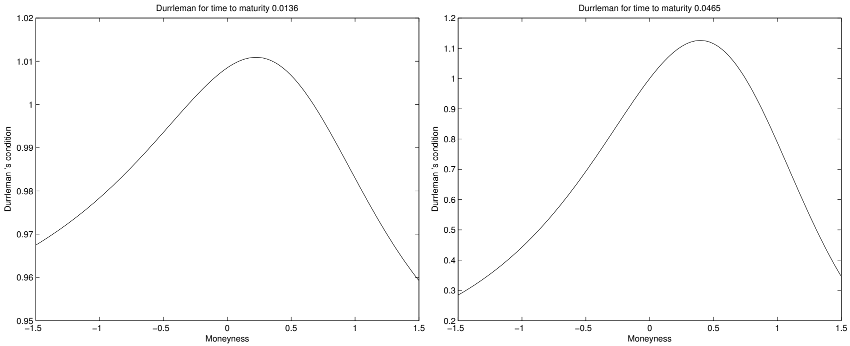
Calendar spread arbitrage

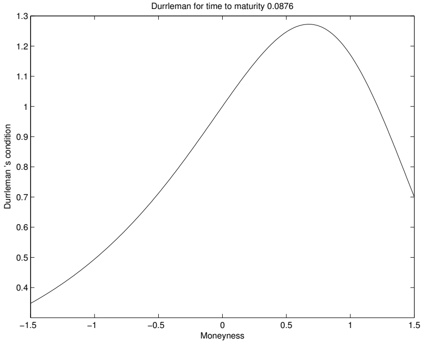
The folowing calibration strategy for each fixed time to maturity makes the volatilty surface free from calendar spread arbitrage

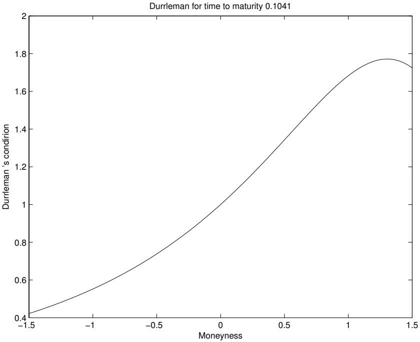
1. $\theta_{2(n)} > \theta_{2(n-1)}$
2. $\theta_{2(n)}\theta_{0(n-1)} + \theta_{2(n-1)}\theta_{0(n)} < \frac{\theta_{1(n)}\theta_{1(n-1)}}{2}$

$\theta_{i(j)}$ is the *i-th* estimated parameter in the optimization for the *j-th* slice.

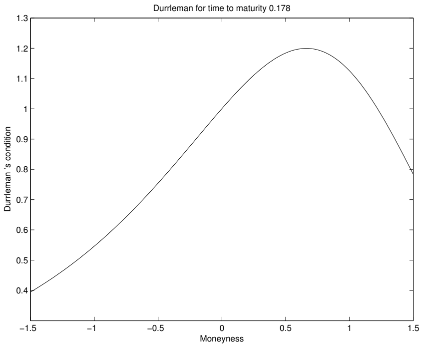
Durrleman 's function

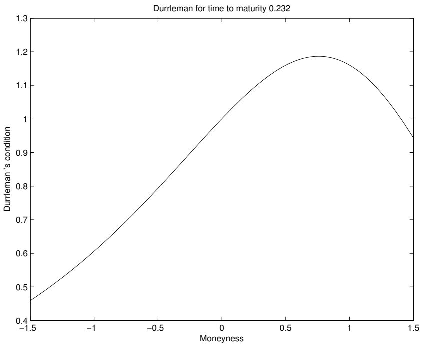



Durrleman for time to maturity 0.0876
Durrleman 's condition
Moneyness
1.3
1.2
1.1
1
0.9
0.8
0.7
0.6
0.5
0.4
−1.5
−1
−0.5
0
0.5
1
1.5


Durrleman for time to maturity 0.1041
Durrleman 's condirion
Moneyness
2
1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
−1.5
−1
−0.5
0
0.5
1
1.5

Durrleman 's function





Calendar spread plot

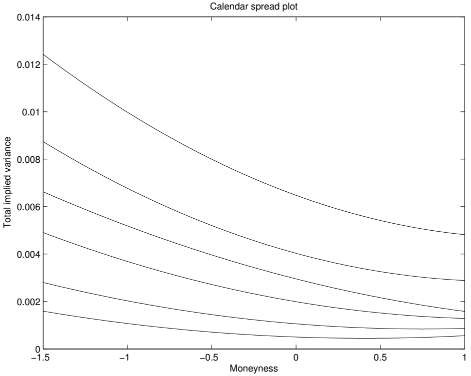


Figure: Plots of total implied variance