

Mathematical & Computational Finance II

Lecture Notes

Computational Aspects of Continuous Time Finance

October 20 2015

Last update: December 4, 2017

1 Barrier Options

Letting $M_t^X = \max_{0 \leq s \leq t} X_s$, for X standard Brownian motion, we can prove the reflection principle for Brownian motion and use this to derive “closed form”-ish¹ formulas for barrier option prices (i.e. up & in, down & out, lookback – anything depending on the max or min in some time interval).

Theorem: The Reflection Principle for Barrier Options. If B is a standard Brownian motion and $M_T^B = \max_{0 \leq t \leq T} B_t$ then we can show that

$$\mathbb{P}(B_T < b, M_T^B > c) = \mathbb{P}(B_T > 2c - b) = \Phi\left(\frac{b - 2c}{\sqrt{T}}\right)$$

where \mathbb{P} is any measure where B is Brownian.

Corollary 1. The joint probability density function of B_T and M_T^B is

$$f(T, b, c) = \frac{2(2c - b)}{T\sqrt{T}} \phi\left(\frac{b - 2c}{\sqrt{T}}\right)$$

The basic idea to the reflection principle is that the first time you reach c you reflect your Brownian path. The strength of this is that once we know the joint PDF (which we have just expressed) then we can glean a lot of information about the max and the min of the system.²

¹That is, closed form up to the normal CDF.

²See “Handbook of Brownian Motion” for all of the formulas you could possibly want to know about Brownian motion (and some you don’t).

2 Calibrating a Volatility Surface

Our goal will be to construct a sufficiently smooth surface for arbitrary strikes K_i and maturities T_i using observed option prices (OTM options) and their implied volatilities.

Consider a response variable Y and predictor variable X and assume that observations are related by

$$Y_i = m(X_i) + \epsilon_i$$

where m is the observation relationship function and ϵ_i is the observation noise. If we assume that $\{\epsilon_i\}_{i=1}^n$ are independent and have mean zero then we may state

$$m(x) = \mathbb{E}[Y|X = x]$$

for a given data set $\{(x_i, y_i)\}_{i=1}^n$. In a local neighbourhood (a small ball) around x containing information about $m(x)$ we estimate $m(x)$ by $\hat{m}(x)$ by locally “averaging” the data around these points. For a given x we say

$$\hat{x}(x) = \frac{1}{n} \sum_{i=1}^n w_{i,n}(x) \cdot y_i$$

where w is some weighting function for each observation. The sequence $\{w_{i,n}\}_{i=1}^n$ denotes the sequence of weights for our observed data. Generally, for $\epsilon > 0$, more weight is given to the values in $(x - \epsilon, x + \epsilon)$. We use a kernel density function $K(x)$ to apply the weights. The kernel density function has properties

1. $K(x) \geq 0$, bounded, continuous, symmetric
2. $\int K(u) du = 1$

Note that this looks like a density probability density function, and so a popular choice is to use the standard normal (Gaussian) density

$$K(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}$$

It is important for our volatility surface to be sufficiently monotone and sufficiently smooth because otherwise there may be (theoretically) arbitrage opportunities for investors. To this extent, note that not all points should appear on the surface since some observations will be anomalous and not line up with theoretical volatilities (i.e. someone paid too much, old data, etc...). In two variables (strike/moneyness and maturity) we use the product of two independent Gaussian densities

$$K(u_1, u_2) = K_1(u_1) \cdot K_2(u_2)$$

To ensure that our surface is smooth enough, that is, to ensure that our kernel is successful, is dependent on the smoothing bandwidth. Define, for $h > 0$

$$K_h(u) = \frac{1}{h} K(u) = \frac{1}{h\sqrt{2\pi}} e^{-\frac{1}{2}(\frac{u}{h})^2}$$

or if we have the joint density of our two variables

$$K_1(u_1) \cdot K_2(u_2) = \frac{1}{h_1 \sqrt{2\pi}} e^{-\frac{1}{2}(\frac{u}{h_2})^2} \cdot \frac{1}{h_2 \sqrt{2\pi}} e^{-\frac{1}{2}(\frac{u}{h_2})^2}$$

We note that

1. h too small leads to an overfit, “too bumpy”, surface
2. h too large leads to an underfit surface that has smoothed away useful information
3. The optimal h can be determined by cross validation (not discussed).

Consider

$$Y = m(X) + \epsilon$$

where

$$\begin{aligned}\mathbb{E}[\epsilon|X = x] &= 0 \\ \mathbb{E}[\epsilon^2|X = x] &= \sigma^2(x)\end{aligned}$$

then we have

$$\begin{aligned}m(x) &= \mathbb{E}[Y|X = x] \\ &= \int y f(y|x) dy \\ &= \int y \frac{f(x, y)}{f(x)} dy \\ &= \frac{\int y f(x, y) dy}{f(x)}\end{aligned}$$

and if we assume the data $\{(x_i, y_i)\}_{i=1}^n$ are i.i.d. we may approximate the integrals by sums to find the Nadaya-Watson estimator for $m(x)$

$$\hat{m}(x) = \frac{\frac{1}{n} \sum_{i=1}^n K_h(x - x_i) \cdot y_i}{\frac{1}{n} \sum_{j=1}^n K_h(x - x_j)}$$

where K_h is the kernel density estimator for $f(x)$. Therefore we see that weights $w_{i,n}(x)$ are

$$\begin{aligned}\hat{m}(x) &= \frac{1}{n} \sum_{i=1}^n w_{i,n}(x) \cdot y_i = \frac{\sum_{i=1}^n K_h(x - x_i) \cdot y_i}{\sum_{j=1}^n K_h(x - x_j)} \\ \implies w_{i,n}(x) &= \frac{K_h(x - x_i)}{\frac{1}{n} \sum_{j=1}^n K_h(x - x_j)}\end{aligned}$$

Note that for a particular observation x_i we have

$$\hat{m}(x) = \frac{K_h(0) \cdot y_i + \sum_{j \neq i}^n K_h(x_i - x_j) \cdot y_j}{\sum_{j=1}^n K_h(x_i - x_j)}$$

and so as $h \rightarrow 0$ we find that³

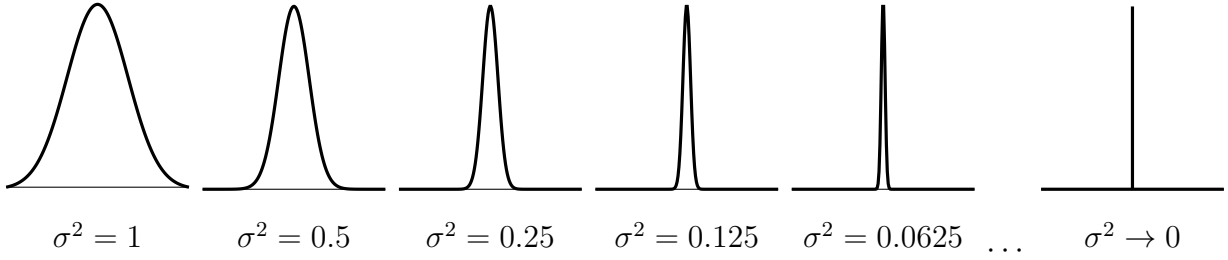
$$\hat{m}(x) \rightarrow \frac{K_h(0)y_i}{K_h(0)} = y_i$$

since if we look at

$$K_h(x) = \frac{1}{h\sqrt{2\pi}} e^{-\frac{1}{2}\frac{x^2}{h^2}}$$

as $h \rightarrow 0$ we get $K_h(x) \rightarrow \delta(x)$ where

$$\delta(x) = \begin{cases} 1 & \text{if } x = 0 \\ 0 & \text{elsewhere} \end{cases}$$



So we see that we get far too much precision (i.e. catastrophically underfit) since this will fit exactly every point to the curve. On the other hand, as $h \rightarrow \infty$ we get $K_h(x) \rightarrow 1$. So we have

$$\lim_{h \rightarrow \infty} \hat{m}(x) = \frac{\frac{1}{n} \sum_{i=1}^n 1 \cdot y_i}{\frac{1}{n} \sum_{j=1}^n 1} = \frac{1}{n} \sum_{i=1}^n y_i = \bar{y}$$

which is just a catastrophically oversmoothed horizontal line at the sample average implied volatility. For implied volatilities the observations are

$$\sigma_{imp}(m_i, T_i)$$

and the response variables are

$$(m_i, T_i)$$

for m_i the moneyness ($\frac{S}{K}$, but there are different parameterizations possible) and time to maturity T_i . So, the 2-D Nadayara-Watson estimator for the implied volatility surface is

$$\hat{\sigma}(m, T) = \frac{\sum_{i=1}^n \sigma_{imp}(m_i, T_i) \cdot g(m - m_i, T - T_i)}{\sum_{j=1}^n g(m - m_i, T - T_i)}$$

where

$$g(x, y) = \frac{1}{2\pi} e^{-\frac{x^2}{2h_1}} \cdot e^{-\frac{y^2}{2h_2}}$$

Note that we have reparameterized h_1^2, h_2^2 with h_1, h_2 since these are just input parameters. We wish to restrict our estimator to a domain

$$(m, t) \in [m_L, m_H] \times [t_L, t_H]$$

³“This just comes from a property of the normal distribution”

where m_L, m_H, t_L, t_H denote some upper and lower bound of moneyness and maturity as great as the min and the max of both variables' observed data. It's important not to extrapolate too far beyond our observed domain because otherwise the space shuttle Challenger explodes. Usually we set $m \in [0.5, 1.5]$, use only OTM price/volatility data, and never use $m = 1$ since this has problematic patterns and behaviour. The decision to select only OTM options is because that ITM options usually have some extra premium attach to them and display unique behaviour. We also avoid using extremely OTM options since these are essentially lottery tickets. That is, severely OTM options should have value zero but do not since the seller must make a premium.

We must also do some data cleaning based on some arbitrage considerations (i.e. removing severe outliers since we predict no arbitrage). In general, we look for

$$\begin{aligned} C_t &\leq (S_t - K)^+ \\ P_t &\leq (K - S_t)^+ \end{aligned}$$

This will be further elaborated upon in an upcoming assignment.

2.1 Example from a Past Exam

Suppose the Black-Scholes implied volatilities of European call and put options on a stock with price $S_0 = 50$ at $t = 0$ are

K	T	Call σ_{imp}	Put σ_{imp}
48	0.25	0.1916	0.1943
49	0.25	0.1982	0.2027
51	0.25	0.2032	0.2145
52	0.25	0.2159	0.2206

Question 1: What does the above data tell us about the validity of the Black-Scholes model?

We see that implied volatilities as computed by the Black-Scholes model are not constant.

Question 2: Outline a method/algorithm for fitting the volatility smile/surface to the above data using a Gaussian kernel $K_h(x)$,

$$K_h(x) = \frac{1}{h\sqrt{2\pi}} e^{-\frac{x^2}{h}}$$

Be as specific as possible and discuss any practical considerations.

Let

$$\hat{m}(x) = \frac{\sum_{i=1}^n K_h(x - x_i) \cdot y_i}{\sum_{j=1}^n K_h(x - x_j)}$$

where (x_i, y_i) are the observed moneyness $x_i = \frac{S}{K_i}$ and the implied volatility

$$y_i = \hat{\sigma}_{imp}(x_i) = \hat{\sigma}_{imp}$$

This gives the fitted smile $(x, m(x))$ for moneyness levels $x \in [x_{low}, x_{high}]$. Note that we should use only OTM implied volatilities to calibrate our model:

K	T	Call σ_{imp}	Put σ_{imp}
48	0.25	–	0.1943
49	0.25	–	0.2027
51	0.25	0.2032	–
52	0.25	0.2159	–

The choice of smooth parameter h is also critical. Selecting h too low will lead to an overfit model that predicts arbitrage and selecting h too high will lead to an underfit model that has smoothed away useful information. Furthermore, we should remove outliers and not extrapolate beyond our observed data.

Question 3: Implement the method outlined in (2) and report the approximation error for $K = 51$ and $h = 0.1$ and $S_0 = 50$.

We know that our kernel density function (in one dimension) looks like

$$K_{0.1}(x - x_i) = \frac{1}{0.1 \cdot \sqrt{2\pi}} e^{-\frac{1}{2} \left(\frac{x - x_i}{0.1} \right)^2}$$

for target moneyness x and observed moneyness x_i . Defining moneyness as $\frac{S}{K_i}$ and letting $x = {}^{50}/_{51}$, we list our data in a table we see

i	K	x_i	σ_{imp}^{OTM}	${}^{50}/_{51} - x_i$	$K_h({}^{50}/_{51} - x_i)$	$K_h({}^{50}/_{51} - x_i) \cdot \sigma_{imp_i}^{OTM}$
1	48	${}^{50}/_{48}$	0.1943	-0.06127	3.3066	0.6425
2	49	${}^{50}/_{49}$	0.2027	-0.04002	3.6825	0.7464
3	51	${}^{50}/_{51}$	0.2032	0.00000	3.9894	0.8106
4	52	${}^{50}/_{52}$	0.2159	0.01885	3.9191	0.8461
				Sum =	14.8976	3.0456

and we find that

$$\hat{\sigma}(x) = \hat{m}({}^{50}/_{51}) = \frac{3.0456}{14.8976} = 0.2044$$

and with approximation error

$$\sigma_{imp_3}^{OTM} - \hat{m}(x) = 0.2032 - 0.2044 = -0.0012$$

Note that our answer is highly dependent on the particular parameterization of moneyness. If we instead defined it instead as $\frac{K}{S}$ then we would see a (slightly?) different answer.

Question 4: Explain how to use the method in (2) to price a European call option expiring in 3 months with strike $K = 49.5$ (no calculations required).

I didn't take notes for this part...