Modeling the Implied Volatility Surface

Abstract
Our study of the volatility surface is twofold. First, we focus on understanding the insufficiency of the Black-Scholes model, and studying the behavior of the volatility surface and volatility skew using AMZN and S&P500 options data. Second, we continue with studying stochastic volatility models to fit dynamics of the volatility skew. We pay a special attention to the GARCH model and Kalman Filtering techniques. Our future work will focus on applying the extended Kalman Filter to a Stochastic Volatility model.

1 Limitation of Black-Scholes Model and Implied Volatility
Volatility is the variation of asset’s returns, which indicates the riskiness of an asset. In options market, theoretical option prices are often calculated under the Black-Scholes Model, which assumes the volatility of the underlying to be known and constant. However, if we invert the Black-Scholes formula given the market option price, the resulting values of volatility vary with the values of Strike (K) and Time to maturity (T). This observation suggests the inconsistency of the Black-Scholes formulation with realized option prices. The empirically calculated volatility is called implied volatility.

To compute implied volatility in R, first we write the Black-Scholes function. Then we write another function called impVol() to calculate the difference between computed options prices and real options prices repeatedly by adjusting the value of σ in the Black-Scholes function till the difference is sufficiently small. We use bisection method to narrow down our selection of σ. We apply impVol() in a for loop to compute implied volatility of the entire dataset automatically.

As shown in Figure 1, plotting the strike price and corresponding implied volatility of options with the same expiration date results in a shape of “smile”. (instead of a horizontal line, as Black-Scholes formula assumes). Implied volatility is a decreasing function of strike price, which indicates that volatility changes are negatively correlated with log returns. The volatility smile occurs because the prices for far out-of-the-money options are bid up higher than the at-the-money strike prices.

Figure 1: Implied Volatility vs Strike for AMZN option expired in December 2017.
Figure 2: IV surface with log strike for S&P500 2009 march Call data
Figure 3: Decay of AMZN 95%-105% skew with respect to maturity.
2 Implied Volatility Surface and Volatility Skew

We then create the volatility surface for S&P500 call options data by computing the collection of implied volatilities across strike prices and time to maturity. Here we use akima and ggplot2 packages to plot the 3-dimensional surface. Again, the Black-Scholes model suggests a flat profile of volatility surface, which conflicts with our empirical observations shown in Figure 2.

The volatility skew is used to characterize the shape of volatility surface. Since empirically, volatility is roughly log-normally distributed, we calculate and plot the “95-105” skew for AMZN call options as a measure of skew, which is the difference between the implied volatilities at strikes of 95% and 105% of the forward price.

In Figure 3, we can observe the following characteristics of volatility skew:

- Short-dated volatilities have higher level of skew than long-dated volatilities.
- Skew with downside implied volatilities are significantly higher than the upside implied volatilities.

3 Stochastic Volatility Model

Since the Black-Scholes model fails to model volatility, we need models to better capture the stylized facts of volatility surface. A widely used type of model is called discrete-time stochastic volatility model, for which the variance of process is randomly distributed. Popular stochastic volatility models include Heston, CEV, SABR, and GARCH. In our study, we focus on evaluating the GARCH model.

3.1 The GARCH Model

The goal of GARCH is to estimate $\sigma$ (conditional standard deviation, which indicates volatility) by estimating $\alpha_i$ and $\beta_j$. It asserts that the predictor of the variance in the next period is a weighted average of the long-run average variance, the variance predicted for this period and the new observation today. Then the model will keep updating new estimation with new information provided. In this case, the weights of the weighted average are $(1-\alpha, \beta, \alpha)$.

The definition of a GARCH(m,s) model is:

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{m} \alpha_i \sigma_{t-i}^2 + \sum_{j=1}^{s} \beta_j \sigma_{t-j}^2,$$

where $\alpha_i > 0, \beta_j > 0, i > 0, j > 0$

$a_t$ is the simple stock return, $\alpha_i$ is the reaction parameter, which indicates the contributions of latest information to conditional variance, $\beta_j$ is the persistence parameter, which indicates the recent level of volatility, $\epsilon_t$ is a sequence of iid random variables with mean 0 and variance 1, $\sigma^2$ is conditional variance, the estimator of volatility, $m$ represents how far back we use the information history, $s$ represents how far back we use the variance history.

Note that in GARCH the only input is stock return. To estimate the parameters of a GARCH model, we use a maximum likelihood estimator by substituting variance of residuals for $\sigma^2$ in normal likelihood and then maximizing with respect to the parameters.

Figure 5: Estimated volatility of the simulated GARCH(1,1) data with $\alpha = 0.85$ and $\beta = 0.1$ vs $\alpha = 0.1$ and $\beta = 0.85$
3.2 Simulation of GARCH(1,1) and Relevant Features
We use fGarch package to simulate and fit data into GARCH model with different levels of $\alpha$ and $\beta$. In Figure 5, we can observe that high $\alpha$ and low $\beta$ will produce more spiky plots (higher volatility of volatility). By contrast, low $\alpha$ and high $\beta$ results in higher level of volatility.

Another important feature of a GARCH model is persistence. Persistence is the time taken for volatility to move halfway back towards its unconditional mean following a deviation from it. In our simulation, $\alpha + \beta = 0.95$. The sum is less than but close to 1, indicating that stocks have a decaying impact on future volatility, the conditional variance is time varying and strongly persistent.

3.3 Filtering for GARCH: Kalman Filter
Kalman Filter is an algorithm that estimates hidden states and variables in a state space model assuming that the observation and state processes are linear and Gaussian. The advantage of the Kalman filter is that it specifies how to update the filter from previous state to current once a new observation is obtained, without having to reprocess the entire data set. Since it is one of the dynamic linear models with state space form, we can possibly apply Kalman filter to model GARCH process for self-adjusting values of volatility.

3.3.1 Kalman filter for linear model & an implementation in CAPM.
To study the features of Kalman Filter in practice, we fit Kalman Filter into capital asset pricing model (CAPM). CAPM describes the relationship between risk and expected return. It says the expected return of a security or portfolio is equal to risk free rate plus a risk premium multiplied by the asset’s systematic risk. We use SSModel, fitSSM, and KFS functions in KFAS R package for modeling and estimation, and compare the resulting $\beta$ parameter with $\beta$ estimated by Ordinary least squares model.

As shown in Figure 6, OLS estimated beta is constant over the period, while Kalman Filter detects the structural changes in systematic risk and adjust to the new information to estimate $\beta$s. The resulting smoothed state $\beta$ is clearly more realistic since there are periods when AMZN stock is less risky than the market, and times when riskier. For example, during 2009 crisis, AMZN was actually less volatile than the market. OLS cannot detect the changes unless calculated manually using rolling window. Notice that Kalman filter also results in a higher average value of $\beta$ compared to OLS estimate, meaning the Kalman filter $\beta$ can be more volatile than OLS estimates.

Figure 6: CAPM $\beta$ estimates from OLS vs Kalman Filter for AMZN stocks return 2007-2014

3.3.2 Extended Kalman Filter for stochastic volatility (future work)
For nonlinear model like GARCH, the original Kalman Filter doesn’t apply. We need to use extended Kalman Filter (EKF), which is a nonlinear version of the Kalman filter. In EKF, the observation and states transition models don’t need to be linear, but instead be differentiable function, so that EKF could linearize the non-linear function around the current estimate. For future work, we will focus on studying EKF and its application on GARCH.
Reference