EXTENDING THE FEATURE SET OF A DATA-DRIVEN ARTIFICIAL NEURAL NETWORK MODEL OF PRICING FINANCIAL OPTIONS

Luis Montesdeoca and Mahesan Niranjan
School of Electronics and Computer Science
University of Southampton, University Road SO17 1BJ
Southampton, UK
Email: {ljm1e14,m.niranjan}@soton.ac.uk

Abstract—Prices of derivative contracts, such as options, traded in the financial markets are expected to have complex relationships to fluctuations in the values of the underlying assets, the time to maturity and type of exercise of the contracts as well as other macroeconomic variables. Hutchinson, Lo and Poggio showed in 1994 that a non-parametric artificial neural network may be trained to approximate this complex functional relationship. Here, we consider this model with additional inputs relevant to the pricing of options and show that the accuracy of approximation may indeed be improved. We consider volume traded, historic volatility, observed interest rates and combinations of these as additional features. In addition to giving empirical results on how the inclusion of these variables helps predicting option prices, we also analyse prediction errors of the different models with volatility and volume traded as inputs, and report an interesting correlation between their contributions.

1. Introduction

Pricing derivative contracts is a challenging problem in financial engineering because contracts mature at some point in the future and there are multiple sources of uncertainty between the current time at which a fair price for the contract needs to be determined and the point at which it may be exercised. The celebrated Black-Scholes model of options pricing makes specific assumptions about a stochastic process model of the underlying asset price and other factors relating to it [1]. These assumptions lead to a solvable differential equation and result in a closed form pricing formula for certain simple derivative instruments. For more complex contracts, where analytical solutions are not possible, Monte Carlo simulation and numerical analysis based methods have been developed. There is significant interest in research literature and wide practical applications of these topics [2].

In this context, the work of Hutchinson, Lo and Poggio in 1994 [3], could be seen as an elegant development from a machine learning or data-driven modelling point of view. Their non-parametric approach discards the restriction of the Black-Scholes model and hence it becomes more adaptive and flexible. The authors showed that a non-parametric artificial neural network, specifically a Radial Basis Functions (RBF) model, can be trained to approximate the complex relationship between the prices of an options contract and the underlying asset. In particular they used only the normalized asset price and the time to maturity of the contract as inputs to the network and further showed that the derivatives of the mapping learned by the network faithfully reproduced the hedge ratio (\(\Delta\), Delta), a widely used parameter in balancing portfolio risk. Neural networks are powerful non-linear approximators, and the RBF architecture itself has found a wide range of applications including speech classification [4], time series prediction [5] and financial engineering [6] among others. RBF is easily deployed in problems that require sequential learning and adaptive model complexity as demonstrated in the resource allocating network [5], [7], [8], and their generalization properties have been analysed in [9], [10].

In addition to the demonstration that options prices may be well-approximated, Hutchinson et al.’s work, which forms the basis of the present study, is notable for a second reason that is of interest in financial engineering. Their work showed that the derivatives of the learned model, which is easily computed for the RBF model analytically, turned out to be good approximations to the hedge ratio, commonly known to practitioners by the Greek letter Delta (\(\Delta\)). This ratio determines the construction of a risk neutral portfolio in which the uncertainty induced by a stochastic process model of asset price changes may be cancelled out. In a later development of Hutchinson et al.’s work, one of us showed that the RBF model, as used in this context, and the Black-Scholes model itself, may be cast as dynamical systems, and the unknowns in the model inferred in a sequential setting using the Extended Kalman Filter (EKF) algorithm [11]. A broader review of the uses of neural networks, with currency options as the application is given in [12].

A further topic in empirical finance is the relationship between the traded volume of an asset and its volatility. Do assets that are traded more show greater price fluctuations, and hence greater uncertainty? While it is tempting to expect such a relationship, there may be no theoretical grounds to reach such a conclusion. Various empirical studies have attempted to test if volume traded and volatility are correlated [13], [14]. For instance, in [15] for futures on the London
International Financial Futures Exchange (Liffe) a positive correlation between these two variables was found. Besides, in [16] found that trading volumes are associated with equity market volatility, in all the markets studied of the BRIC (Brasilia, Russian, India and China) countries.

An underlying theoretical premise, known as the Mixture of Distribution Hypothesis (MDH), introduced by Clark [17], suggests that daily trading volume and price changes are driven by the same flow of information. Starting from this there are empirical studies that have attempted to explore this relationship. For example, Wen-Cheng et al. [18] suggest a bidirectional relationship between the two using a bivariate vector autoregressive methodology, and Jain [19] suggests a strong contemporaneous relation between trading volume and volatility with hourly values. However, Karpoff [20] reports an asymmetric relationship, which is also supported by the work in [21]. Similar explorations have been carried out on the Korean and New York Stock Exchange data [22], [23]. Nevertheless, in various commodity future contracts analyzed in the Chinese market, the correlation between return and trading volume are close to zero [24] and on intra-day trading data on S&P 500 index a negative correlation is reported [25].

In this work, we extend Hutchinson et al.’s work by asking the question if expanding the feature set to include additional variables of interest help in improving their data driven model of options pricing. Specifically, we include combinations of historic volatility, volume traded and interest rates as inputs in addition to the normalized asset price and time to maturity, as illustrated in Figure 1. Empirical work we carried out, on a range of data with much wider scope than the original work, shows this to be the case. We follow this up with an analysis of how much the volume traded and volatility of asset help in predicting options prices and demonstrate an intriguing correlation between their relative contributions.

2. Model & Data

In this section, we describe the architecture of the Radial Basis Functions (RBF) model as introduced in [3], our extensions to the feature sets and the data used in training the networks. We use substantially more data than [3], who restricted their analysis to options drawn on the S&P500 Index only.

Let the input data vector to the RBF model given by the vector
\[
x = \begin{bmatrix} S \\ X \\ T - t \end{bmatrix}^T, \tag{1}
\]
where \(S\) denotes the underlying asset price, \(X\), the strike price of the contract and \(T - t\) is the time to maturity (the difference between the time of maturity, \(T\), of the contract and the present time \(t\)). With this vector of features as input, each of the basis functions in the RBF model for predicting options prices is written as
\[
\phi_k = \sqrt{\left( x - m_k \right)^T \sum_k^{-1} \left( x - m_k \right) + b_k}, \tag{2}
\]
where, \(\phi_k\) denotes the response of a nonlinear basis function which is parameterized by local mean \(m_k\) and covariance matrix \(\Sigma_k\) and \(b_k\), a local bias term.

The basis function locations and local covariance matrices are estimated by fitting a Gaussian Mixture Model (GMM) to the distribution of input data, making the model sensitive to its local density. For simplicity, in our implementations, we set the bias terms \(b_k\) to zero. In addition to these nonlinear terms, the model includes a linear part as well. Thus least squares problem to solve is shown by the simultaneous equations:
\[
\begin{bmatrix} \phi_{11} & \ldots & \phi_{14} \\ \phi_{21} & \ldots & \phi_{24} \\ \vdots & \ldots & \vdots \\ \phi_{n1} & \ldots & \phi_{n4} \end{bmatrix} \begin{bmatrix} x_{11} \\ x_{21} \\ \vdots \\ x_{n1} \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_4 \\ \vdots \\ \lambda_n \end{bmatrix}
\]
\[
= \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix} \tag{3}
\]
where \(n\) is the number of observations and \(c_j, j = 1, \ldots n\) are the output call option prices normalized by their strike prices. The vector of unknowns \([ \lambda_1 \ldots \lambda_4 w_1 w_2 w_0]^T\) is estimated by linear least squares, which is the main computational advantage of using a fixed non-linear, linear in unknowns RBF model.

A pseudo-inverse solution to the problem is often used
\[
w = (P^TP)^{-1}P^T \Phi, \tag{4}
\]
where \(P\) is the so called design matrix and \(\Phi\), the vector of outputs. Often, for reasons of numerical ill-conditioning and to avoid over-fitting by the model, a regularization term in the form of a diagonal matrix is added before the inversion of \(P^TP\):
\[
w = (P^TP + \gamma I)^{-1}P^T \Phi, \tag{5}
\]
where $\gamma$ controls how much regularization is applied.

Once the model is trained, the resulting output is given by,

$$\hat{c} = \Phi \lambda + w^T x + w_0 \quad (6)$$

Following the work of Hutchinson et al. we have defined the above model with four basis functions. In our own work, we tested the effect of the choice of the number of basis functions, hence the model complexity, and found four to be a reasonable number. We also confirmed this by running the RBF model with the Akaike information criterion (AIC) [26] as the method of model selection. Beyond that the model showed clear signs of overfitting as shown in Figure 2.

For empirical evaluation, we used the daily prices of 21 call and 11 put options written on the Financial Times Stock Exchange (FTSE) index and nine call and seven put options of the popular software companies Apple and Microsoft. Additionally, we considered minute-by-minute intraday call and put options prices on a contract on Apple with a strike price of 95 and 100, maturing in September 2015 and June 2016 respectively. Table 1 lists the range of options considered and includes their strike prices and dates of maturity.

For selection of training data, for any of the options we took a temporal window of 40% of the data as the training set and evaluated the model performance on the next point in time. We moved this window one sample at a time and repeated the training and testing. Thus all results we quote are based on single sample unseen data. The reason for this choice of window, rather than a randomized training/test partition as often used in machine learning is that the financial data is expected to have temporal structure and in any practical application, one is likely to apply a trained model in the next point in time. We computed volatility over a window of 25% of the data immediately preceding a point of analysis.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>C6800</td>
<td>09-13</td>
<td>171</td>
<td>C6600</td>
</tr>
<tr>
<td>C6700</td>
<td>09-13</td>
<td>290</td>
<td>C6700</td>
</tr>
<tr>
<td>C6800</td>
<td>09-13</td>
<td>171</td>
<td>C6800</td>
</tr>
<tr>
<td>C6700</td>
<td>09-14</td>
<td>171</td>
<td>C6900</td>
</tr>
<tr>
<td>C6750</td>
<td>09-14</td>
<td>525</td>
<td>C7000</td>
</tr>
<tr>
<td>C6800</td>
<td>09-14</td>
<td>291</td>
<td>P6000</td>
</tr>
<tr>
<td>C6900</td>
<td>09-14</td>
<td>288</td>
<td>P6300</td>
</tr>
<tr>
<td>C7000</td>
<td>09-14</td>
<td>288</td>
<td>P6500</td>
</tr>
<tr>
<td>C7100</td>
<td>09-14</td>
<td>299</td>
<td>P6700</td>
</tr>
<tr>
<td>C6600</td>
<td>09-14</td>
<td>462</td>
<td>P6900</td>
</tr>
<tr>
<td>C6800</td>
<td>09-14</td>
<td>462</td>
<td>P6900</td>
</tr>
<tr>
<td>C6900</td>
<td>09-14</td>
<td>462</td>
<td>P7000</td>
</tr>
<tr>
<td>C7000</td>
<td>09-14</td>
<td>462</td>
<td>P7000</td>
</tr>
<tr>
<td>C7200</td>
<td>09-14</td>
<td>462</td>
<td>P7000</td>
</tr>
<tr>
<td>C7400</td>
<td>09-14</td>
<td>342</td>
<td>P6800</td>
</tr>
<tr>
<td>CAPL50</td>
<td>07-15</td>
<td>239</td>
<td>CMSF39</td>
</tr>
<tr>
<td>CAPL85</td>
<td>07-15</td>
<td>239</td>
<td>PAPL120</td>
</tr>
<tr>
<td>CAPL90</td>
<td>07-15</td>
<td>239</td>
<td>PAPL130</td>
</tr>
<tr>
<td>CAPL95</td>
<td>07-15</td>
<td>239</td>
<td>PAPL135</td>
</tr>
<tr>
<td>CAPL105</td>
<td>07-15</td>
<td>236</td>
<td>PAPL140</td>
</tr>
<tr>
<td>CAPL110</td>
<td>07-15</td>
<td>236</td>
<td>PAPL150</td>
</tr>
<tr>
<td>CAPL115</td>
<td>07-15</td>
<td>236</td>
<td>PMSF45</td>
</tr>
<tr>
<td>CAPL120</td>
<td>07-15</td>
<td>236</td>
<td>PMSF47</td>
</tr>
</tbody>
</table>

3. Empirical Results

3.1. Prediction performance with additional features

Evaluating the options listed in the Table 1 demonstrate that Hutchinson et al.’s model with additional inputs relevant to the pricing of options may have enhanced the accuracy of approximation. This values are shown in Table 2 for call and put options. An example of this improvement is illustrated in Figure 3.
TABLE 2. AVERAGES OF MEAN SQUARED MODELLING ERRORS OF THE VARIOUS MODELS ON THE DIFFERENT OPTION CONTRACTS. THE DISTRIBUTION OF MEAN SQUARED ERRORS IS SHOWN IN FIGURE 4. EACH VALUE IN THE TABLE SHOULD BE SCALED BY $10^{-6}$.

<table>
<thead>
<tr>
<th></th>
<th>Hutch.</th>
<th>Vol.</th>
<th>Sig</th>
<th>Sig+Vol</th>
<th>LR.</th>
<th>IR+Vol</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>5.19</td>
<td>4.27</td>
<td>5.51</td>
<td>4.79</td>
<td>4.70</td>
<td>4.21</td>
</tr>
<tr>
<td>P</td>
<td>9.13</td>
<td>8.78</td>
<td>9.51</td>
<td>9.74</td>
<td>8.38</td>
<td>8.93</td>
</tr>
</tbody>
</table>

Figure 4. Prediction performance on FTSE index call options for the various models considered. Overall, the largest improvement in prediction accuracies is obtained when volume traded and interest rates are included as additional covariates.

Figure 5 shows performance of RBF predictors on the two equity options. Here, with volatility as additional input we see only a marginal improvement in performance whereas all other models give significant reduction in errors when compared to Hutchinson et al.’s model.

Figure 6 shows performance of RBF predictors considering minute-by-minute intraday call and put equity options. Here, we plot the squared error at every point in time (minute) as boxplots of the intraday call option. The corresponding mean squared errors are shown in Table 3.

TABLE 3. AVERAGE PREDICTION PERFORMANCES OF VARIOUS MODELS ON INTRA-DAY CALL AND PUT OPTIONS CONTRACT. EACH VALUE IN THE TABLE SHOULD BE SCALED BY $10^{-5}$.

<table>
<thead>
<tr>
<th></th>
<th>Hutchinson</th>
<th>Volume</th>
<th>Sigma</th>
<th>Sigma+Volume</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>2.97</td>
<td>2.19</td>
<td>2.64</td>
<td>2.76</td>
</tr>
<tr>
<td>P</td>
<td>2.72</td>
<td>2.57</td>
<td>2.83</td>
<td>2.29</td>
</tr>
</tbody>
</table>

3.2. Volume Traded and Volatility

Further we used these models as a tool to explore the relationships between volatility and trading volume. The measurement of this correlation is made by the mean squared errors of the predictive models with volume traded and volatility as additional inputs. Figure 7 (left boxplot) shows the distribution of correlations in mean squared errors of model fitting with volume traded and volatility as additional inputs on all 48 options which is 0.675. In the right boxplot in Figure 7 shows the correlation between the corresponding volume trading values and volatility values.

The observations analysed for this empirical study are separated into equity options and FTSE 100 index options. Figure 8 shows a substantial correlation between volume traded and volatility in predicting options prices in the FTSE 100 index call options. Evaluating the call options found a Pearson coefficient value of 0.663 in the Mean Square Errors on the RBF with volume and volatility as additional inputs. A similar distribution was observed for put options with a correlation of 0.678 pricing put derivative options.

Analysing the equity call and put options, Figure 9 shows the Pearson correlation coefficient between Mean Square Errors on the RBF with volume and volatility as additional inputs. The mean on the correlation of each call
Figure 7. The Pearson correlation values from the mean square errors of the RBF models with volatility and trading volume on all options tested. It demonstrates that the mean on the correlation of each option analysed is over 0.67 giving a substantial correlation. The relationship between the corresponding values of trading volume and volatility does not show a correlation.

Figure 8. Correlation on the contribution of RBF errors on predicting each call FTSE options when volume and volatility are included as additional inputs. This correlation value is 0.663.

Figure 9. Correlation between Mean Square Error values of the RBF models that include trading volume and volatility on put(a) and call(b) equity options.

Figure 10. Contribution to prediction error from volatility and volume traded on minute-by-minute intra-day data on equity call option, showing a high correlation of 0.72.

4. Conclusion

In this work, we study a non-parametric neural network model that quantifies the complex relationship between a class of financial instruments known as options and that of the underlying asset on which the contract is drawn. Whereas previous work introducing this model uses the asset price and the time to maturity of the contract as its only inputs, we have demonstrated that the inclusion of additional features relating to the contract, namely the volatility, volume of the underlying asset traded and the risk-free interest rate help in improving the accuracy with which the market value of the contract may be predicted. Our empirical results, carried out on index options, two equity options and an intra-day contract covers a substantially wider range than previous authors have considered. Hence the results quoted in this study are more robust.

The results lead to the exploration of an intriguing relationship between the volatility of an asset and the amount of volume traded, a topic that has attracted healthy discus-
practical application of VC formulation and inference using extended Kalman filter has been shown to be possible [11]. Hence our current work focuses on extending the models with the extended features to a dynamical setting.

References