Incorporating Prior Financial Domain Knowledge into Neural Networks for Implied Volatility Surface Prediction

Yu Zheng[1,2], Yongxin Yang[3,4], Bowei Chen[5,*]

zhengyu@swufe.edu.cn | yongxin.yang@ed.ac.uk | bowei.chen@glasgow.ac.uk

1. Southwestern University of Finance and Economics 2. Inboc Technologies Ltd.
3. University of Edinburgh 4. ArrayStream Technologies Ltd. 5. University of Glasgow

KDD2021

Introduction

In this paper we develop a novel neural network model for predicting implied volatility surface. Prior financial domain knowledge is taken into account. A new activation function that incorporates volatility smile is proposed, which is used for the hidden nodes that process the underlying asset price. In addition, financial conditions, such as the absence of arbitrage, the boundaries and the asymptotic slope, are embedded into the loss function. This is one of the very first studies which discuss a methodological framework that incorporates prior financial domain knowledge into neural network architecture design and model training. The proposed model outperforms the benchmarked models with the option data on the S&P 500 index over 20 years. More importantly, the domain knowledge is satisfied empirically, showing the model is consistent with the existing financial theories and conditions related to implied volatility surface.

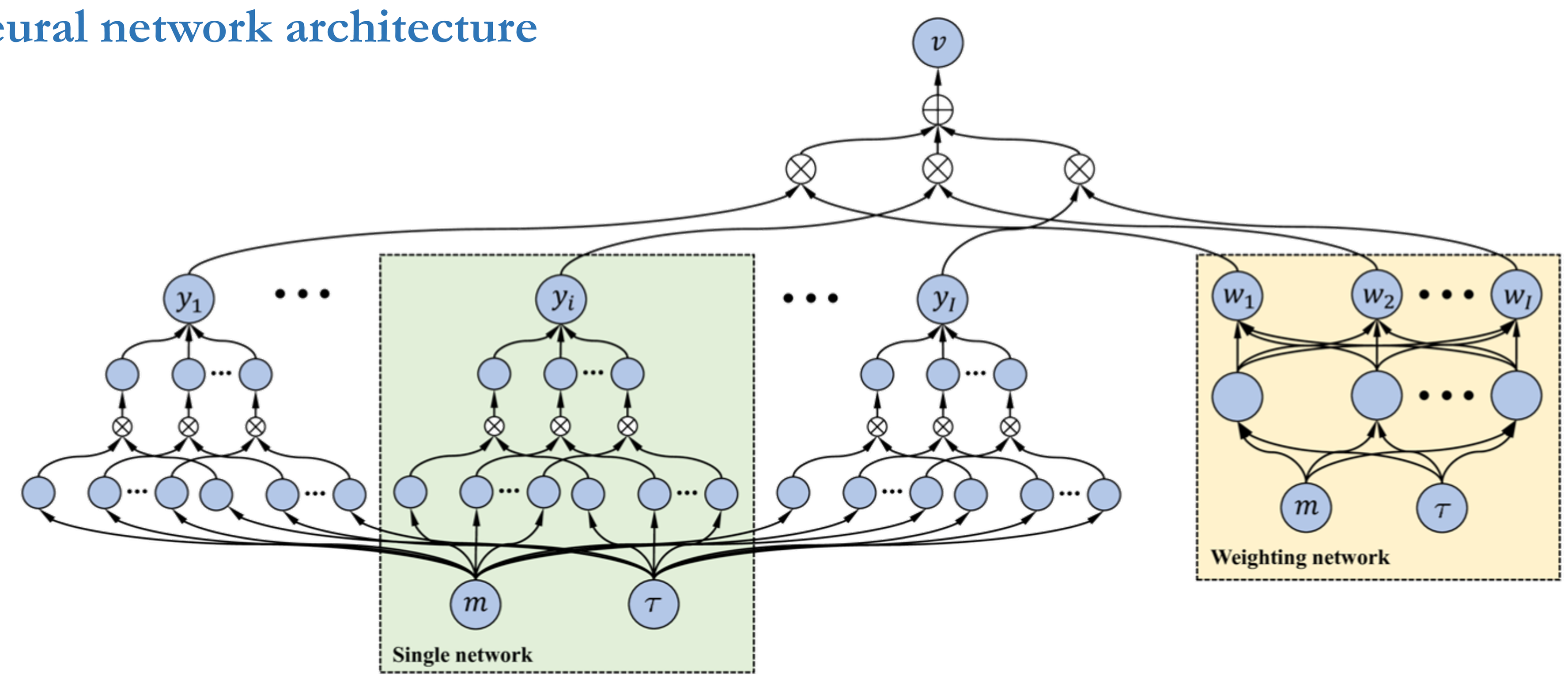
Implied volatility surface

The **implied volatility** of an option is defined as the inverse problem of option pricing, mapping from the option price in the current market to a single value. When it is plotted against the option price strike price and the time to maturity, it is called the **implied volatility surface**.

$$\begin{aligned} C(X_t,t) &= N(d_1)X_t - N(d_2)Ke^{-r(T-t)}, \\ d_1 &= \frac{1}{\sigma\sqrt{T-t}}\left[\ln\left(\frac{X_t}{K}\right) + \left(r+\frac{\sigma^2}{2}\right)(T-t)\right], \\ d_2 &= d_1 - \sigma\sqrt{T-t}, \end{aligned}$$

Black, F., & Scholes, M. (1973) The pricing of options and corporate liabilities. *Journal of Political Economy*, 81, 637–654.

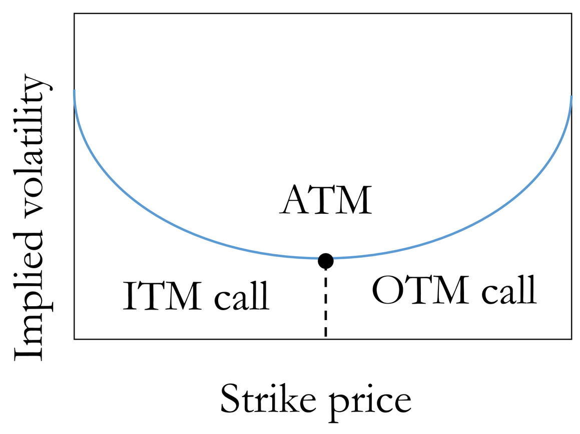
Neural network architecture



Empirical evidence volatility smile

Given time to maturity, if the implied volatility is plotted against the strike price, it creates a line that slopes upward on either end, looking like a "smile".

We propose a smile activation function for the nodes that correspond to the underlying m



$$\phi(z) = \sqrt{z\tanh(z+\frac{1}{2}) + \tanh(-\frac{1}{2}z+\epsilon)},$$

Embedding constraints into optimisation

$$\min \ell = \ell_0 + \gamma\ell_1 + \delta\ell_2 + \eta\ell_3 + \rho\ell_4 + \omega\ell_5,$$

$$\ell_0 = \alpha\underbrace{\left(\frac{1}{N}\sum_{n=1}^{N}(\log(v_n) - \log(\hat{v}_n))^2\right)}_{\text{MSLE}} + \beta\underbrace{\left(\frac{1}{N}\sum_{n=1}^{N}(\frac{v_n-\hat{v}_n}{v_n})^2\right)}_{\text{MSPE}},$$

$$\ell_1 = \sum_{p=1}^{P}\sum_{q=1}^{Q}\max(0, -a(m_p,\tau_q)), \qquad \triangleright\, a(m,\tau) := v(m,\tau) + 2\tau\partial_\tau v(m,\tau)$$

$$\ell_2 = \sum_{p=1}^{P}\sum_{q=1}^{Q}\max(0, -b(m_p,\tau_q)),$$

$$\triangleright\, b(m,\tau) := (1 - \tfrac{m\partial_m v(m,\tau)}{v(m,\tau)})^2 - \tfrac{(v(m,\tau)\tau\partial_m v(m,\tau))^2}{4} + \tau v(m,\tau)\partial_{mm}v(m,\tau)$$

$$\ell_3 = \sum_{p_1=1}^{P_1}\sum_{q=1}^{Q}\max(0, -c_1(m_{p_1},\tau_q)) + \sum_{p_2=1}^{P_2}\sum_{q=1}^{Q}\max(0, -c_2(m_{p_2},\tau_q)),$$

$$\triangleright\, c_1(m,\tau) := N(d_-(m,\tau)) - \sqrt{\tau}\partial_m v(m,\tau)n(d_-(m,\tau))$$

$$\triangleright\, c_2(m,\tau) := N(-d_-(m,\tau)) + \sqrt{\tau}\partial_m v(m,\tau)n(d_-(m,\tau))$$

$$\ell_4 = \sum_{p=1}^{P}\sum_{q=1}^{Q}\max(0, -(g(m_p,\tau_q) - \epsilon)), \qquad \triangleright\, g(m,\tau) := 2|m| - v^2(m,\tau)\tau$$

$$\ell_5 = \begin{cases} ||\bar{W}||_F^2 + ||\tilde{W}||_F^2 + ||\hat{W}||_F^2, & \text{for the single model,} \\ \sum_{i=1}^{I}||\bar{W}^{(i)}||_F^2 + \sum_{i=1}^{I}||\tilde{W}^{(i)}||_F^2 + \sum_{i=1}^{I}||\hat{W}^{(i)}||_F^2 + ||\dot{W}||_F^2 + ||\ddot{W}||_F^2, & \text{for the multimodel.} \end{cases}$$

Financial conditions

Absence of arbitrage conditions

Gulisashvili, A. (2012) Analytically Tractable Stochastic Stock Price Models. Springer.

- Positivity
- Twice differentiable
- Monotonicity
- Absence of the butterfly arbitrage
- Limit condition

Boundaries conditions

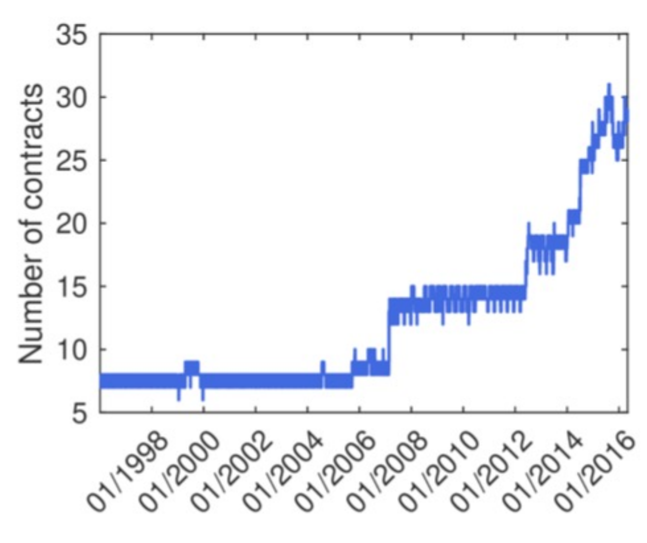
Carr, P., & Wu, L. (2007) Stochastic skew in currency options. *Journal of Financial Economics*, 86, 213–247.

Asymptotic slope condition

Lee, R. (2004) The moment formula for implied volatility at extreme strikes. *Mathematical Finance.* 14(3), 469–480.

Experiments

Data (S&P 500 options)



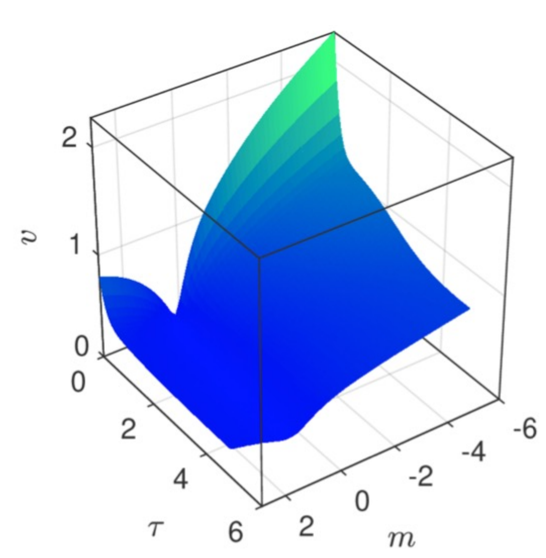
Implied volatility

Model	Training		Test	
	Mean	STD	Mean	STD
Multi	1.74	0.50	3.34	2.18
Multi†	1.76	0.50	3.35	2.17
Single	2.15	0.67	3.60	2.12
Single†	1.82	0.52	3.38	2.16
Vanilla	3.21	0.98	4.46	2.07
Vanilla†	2.87	0.80	4.18	2.04
SSVI	2.59	0.85	3.73	2.18

Option price

Model	Training		Test	
	Mean	STD	Mean	STD
Multi	5.97	1.86	10.64	6.72
Multi†	6.03	1.86	10.67	6.70
Single	7.38	2.57	11.64	6.68
Single†	6.20	1.91	10.77	6.67
Vanilla	11.31	3.57	14.61	6.42
Vanilla†	10.53	3.34	14.17	6.60
SSVI	8.71	2.72	12.74	6.74

Predicted implied volatility surface



Conditions check

Model	Monotonicity	Absence of butterfly arbitrage	Left boundary	Right boundary	Asymptotic slope
Multi	0.00%	7.02e-6%	0.00%	0.00%	0.00%
Multi†	1.28%	4.87%	0.00%	14.06%	0.00%
Single	0.00%	5.56e-3 %	0.00%	0.05%	0.00%
Single†	0.00%	14.88%	0.95%	5.16%	0.00%
Vanilla	3.75e-3%	1.53e-2%	4.07e-3%	0.00%	0.00%
Vanilla†	5.32%	5.72%	14.63%	0.00%	0.54%