Fair and Explainable Heavy-tailed Solutions of Option Prices through Reinforcement, Deep, and EM Learnings

Chansoo Kim
eau@econ.kist.re.kr
Computational Science Center & ESRI
Korea Institute of Science and Technology

Department of Economics & SIRFE
Seoul National University
Seoul, Republic of Korea

ByoungSeon Choi
bschoi12@snu.ac.kr
Department of Economics & SIRFE
Seoul National University
Seoul, Republic of Korea

We compute option prices of the Korean Stock Prices Index (KOSPI) through three popular data analytics approaches: reinforcement learning (Q-learning), deep learning and expectation maximization algorithm. The computed results are compared with the generalized Black-Scholes option prices based on the underlying with leptokurtic Gram-Charlier A series distribution. These results well explain the non-Gaussian and heavy-tailed behavior of the KOSPI in contrast to the original Black-Scholes formula. Among the three machine learning computations, the reinforcement learning fits best with the analytic solution. The closeness between the reinforcement learning result and heavy-tailed analytic solution in option pricing explains the machine learning result more inductively.

KEYWORDS

Reinforcement Learning, Deep Learning, Expectation Maximization, Option Pricing, Black-Scholes price, Gram-Charlier A series, KOSPI

ACM Reference Format:
Chansoo Kim and ByoungSeon Choi. 2019. Fair and Explainable Heavy-tailed Solutions of Option Prices through Reinforcement, Deep, and EM Learnings. In *The 1st Workshop on Deep Reinforcement Learning for Knowledge Discovery (DRL4KDD '19), August 5, 2019, Anchorage, AK, USA.* ACM, New York, NY, USA, 4 pages.

ACM ISBN 978-x-xxxx-xxxx-x/YY/MM…$15.00

INTRODUCTION

We analyze index option prices using various machine learning approaches such as reinforcement learning (Q-learning), deep (neural network) learning, and expectation maximization (EM) algorithm. Korea Composite Stock Price Index 200 (KOSPI 200) option prices during the fiscal years from 1994 to 2018 are used for our analysis, because it shows heavy-tailed behavior such as kurtosis as well as volatility puzzle well. KOSPI is the index of all common stocks traded on the Stock Market Division of the Korea Exchange. KOSPI 200, which is composed of 200 big companies in the Market, is one of the representative indexes in Republic of Korea. It is equivalent to the S&P 500 in the United States. Fig. 1 depicts its time-varying movement.

The well-known Black Scholes partial differential equation (PDE) is the call option price dynamics over time:

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV = 0, \tag{1}$$

where trading underlying assets eliminate the risks in the market. In Eq. (1), V is the option's price as a function of stock price S and time t, r is the risk-free interest rate, and σ means the volatility of the stock.

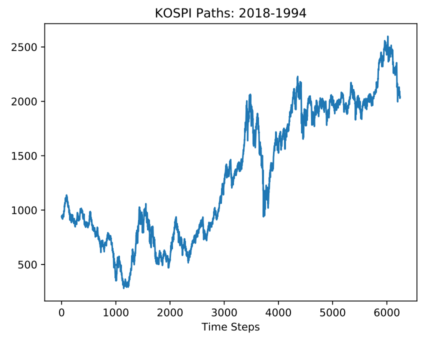

Figure 1: KOSPI Index during a period from 1994 to 2018. Blue line depicts the time-series of KOSPI data.

By solving Eq. (1), we get the Black-Scholes formula:

$$w^{BS}(x_t, t) = N(d_1)\, x_t - N(d_2)\, K e^{-r(T-t)}, \tag{2}$$
$$d_1 = \frac{1}{\sigma\sqrt{T-t}}\left[\ln\left(\frac{S_t}{K}\right) + \left(r + \frac{\sigma^2}{2}\right)(T-t)\right],$$
$$d_2 = d_1 - \sigma\sqrt{T-t}.$$

This model requires the Black-Schols environment, which requires the underlying are impractically perfect. For example it assumes the normality of the return rate. However, many empirical results show that a return rate of stock price is heavy-tailed since Fama [4] and Mandelbrot [8].

To derive a general equilibrium of the option price, we should consider all the assets in the market. It is practically impossible, and it causes the curse-of-dimensionality problem. Moreover the normality assumption needs only two parameters. But, these two parameters cannot represent the heavy-tailed behavior the asset return. Recently it is shown that there exist infinitely many solutions to the boundary problem consisting Black-Schloes PDE and the call option payoff as the terminal condition [1]. The analytic solutions include the ones based on heavy-tailed Gram-Charlier A type distributions, which incorporate the market data more than Gaussian assumption, as well as the famous Black-Schloes formula as a special case.

In this work, we apply three tangible machine learning techniques to evaluate its prices based on the heavy-tailed property. Among the machine learning results, we choose the most suitable to the most appropriate heavy-tailed analytical solution.

Our novelty comes from comparing the black-box type machine learning prices with the analytical one. It gives us some inductive reasonings to explain the learning process.

DATA ANALYTICS: MACHINE LEARNING

Among recent popular machine learning techniques, we choose three important ones for the KOSPI 200 index option pricing: reinforcement learning, deep learning and EM algorithm. We apply KOSPI dataset as shown in Fig. 1: its price dynamics from 1994 to 2018.

Using recent work on Black Scholes world with the reinforcement learning [5], the Index Option in Korean market is priced. It solves the same problem, Eq. (1), in a way of dynamic programming. In other words, it learns for an optimal policy based on samples, while Black-Scholes PDE ignores risk in options. The reinforcement learning incorporates Black-Scholes pricing concept but does not assume famous 'no arbitrage conditions' and exact PDE model, which means existence of known solution [5].

Based on the Bellman equation with the sample target value, which works as a feedback, it tries to solve the followings recursively:

$$\text{action-value: } Q_t^*(X_t, a_t^*) = \gamma \mathbb{E}_t\left[Q_{t+1}^*(X_{t+1}, a_{t+1}^*) - \lambda\gamma\hat{\Pi}_{t+1}^2 + \lambda\gamma\left(a_t^*(X_t)\right)^2\left(\Delta\hat{S}_t\right)^2\right] (t = 0, \ldots, T-1)$$
$$\text{optimal action: } a_t^*(X_t) = \frac{\mathbb{E}_t\left[\Delta\hat{S}_t\hat{\Pi}_{t+1} + \frac{1}{2\gamma\lambda}\Delta S_t\right]}{\mathbb{E}_t\left[\left(\Delta\hat{S}_t\right)^2\right]}$$
$$\text{pricing: } C_t(S_t, ask) = -Q_t(S_t, a_t^*), \tag{3}$$

where notations are defined as in Eq. (1): X_t is underlying (KOSPI index) price at time t and $a_t(X_t)$ is optimal action for the given price X_t. Call option price is determined by Q values as a function of those two at t. Hyperparameters of λ and γ are determined regarding best distribution fit to the underlying.

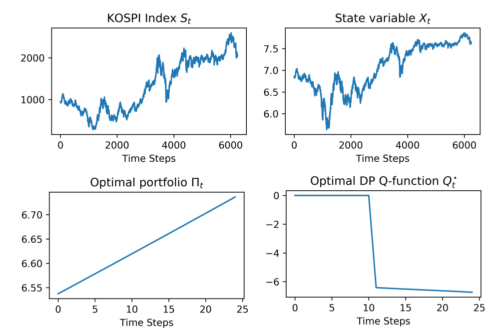

Figure 2: KOSPI 200 option price computed from reinforcement learning. This shows parameters obtained from real KOSPI market data such as Index. As the Q-learning result, we obtain its price of 1,820 and bottom panels show the learning process.

Following Han et al. [6], we apply deep learning technique to solve the Black-Scholes PDE. As how KOSPI 200 Index option is composed, this model consider the fair price of a European claim based on these assets with each strike prices. Then we solve the Black-Scholes option pricing equation for Korean market case with specific rate $r = 0.032$, which is extracted from Fig. 1.

$$\frac{\partial w}{\partial t}(t,x) + \overline{\mu}x \cdot \nabla w(t,x) + \frac{\overline{\sigma}^2}{2}\sum_{i=1}^{d}|x_i|^2\frac{\partial^2 w}{\partial x_i^2}(t,x) - (1-\delta)Q(w(t,x))w(t,x) - Rw(t,x) = 0 \tag{4}$$

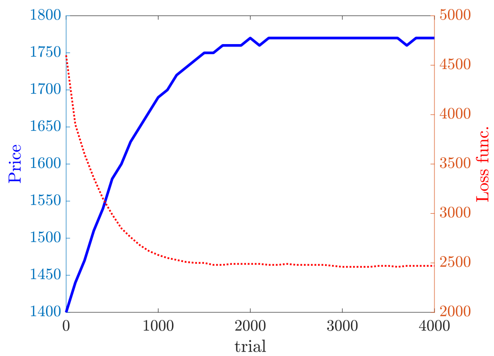

Figure 3: The option price obtained through deep learning. We directly solve the Black-Scholes PDE using deep learning technique. Blue line depicts the price with increasing learning trials. Red dotted line means loss function value for each trial.

Third, we apply expectation maximization learning to our data. It fits the KOSPI 200 underlying return with three Gaussian distribution to obtain a plausible heavy-tail distribution. Three Gaussian distributions fit the distribution, as seen by loglikelihood in Fig. 4.

$$EM = \alpha_{-1}\phi(\mu_{-1}, \sigma_{-1}) + \alpha_0\phi(\mu_0, \sigma_0) + \alpha_{+1}\phi(\mu_{+1}, \sigma_{+1}), \tag{5}$$

where $\phi(\cdot)$ means gaussian density functions.

The EM algorithm works as determining unknown parameters θ through the maximum likelihood estimate (MLE) with the observed KOSPI rate X and its latent data Z.

$$L(\theta; X) = p(X|\theta) = \int p(X, Z|\theta)dZ \tag{6}$$

Under the gaussian distribution, the EM algorithm try to obtain MLE of the marginal likelihood by iteratively applying E and M steps,

$$\begin{aligned} &\text{E step: } Q\left(\theta|\theta^{(t)}\right) = E_{Z|X,\theta^{(t)}}[\log L(\theta; X, Z)] \\ &\text{M step: } \theta^{(t+1)} = \arg\max_{\theta} Q\left(\theta|\theta^{(t)}\right). \end{aligned} \tag{7}$$

In E step, the expected value of the log likelihood function of θ, $Q\left(\theta|\theta^{(t)}\right)$ is computed, and in M step, we find the parameters, which makes this value.

As computed result through Eq. (7), we obtain parameters for the Eq. (5). The first term has $\mu_{-1} = -1.10E-3$ and $\sigma_{-1} = 4.63E-5$ with weight $\alpha_{-1} = 0.457$. The second and third have $\mu_0 = 3.69E-4$ and $\sigma_0 = 2.15E-4$ with weight $\alpha_0 = 0.376$ and $\mu_{+1} = 2.30E-3$ and $\sigma_{+1} = 9.98E-4$ with weight $\alpha_{+1} = 0.167$, relatively.

EM learning works as an important augmenting computation, in the sense that it gives a set of distribution parameters to better use the Black Scholes PDE, Eq. (1). This result can be directly compared to the explainable solution suggested in Eq. (9) below.

EXPLAINABLE SOLUTION: A HEAVY-TAILED ONE

According to Corrado & Su and Jondeau et al., [2, 3, 7], we can model the underlying distribution, which consists of a financial

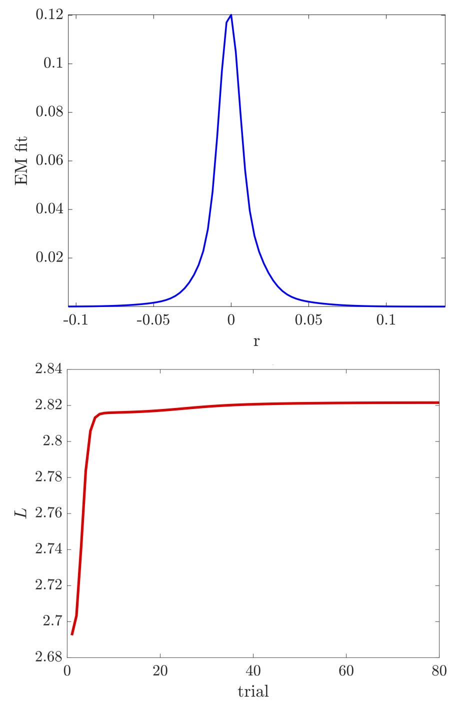

Figure 4: EM fit and its loglikelihood of the KOSPI 200 return distribution by EM algorithm. Upper panel shows EM learning result shown by blue line, and the red line in the lower panel depicts its loglikelihood suggesting best result.

option, using Gram-Charlier A series with kurtosis and leptokurtic.

$$\begin{aligned} q^{\boldsymbol{\eta}}\left(x; h_t, v^2\tau\right) \equiv &\left\{1 + \eta_2\frac{(x-h_t)^2 - v^2\tau}{v^4\tau^2}\right. \\ &\left. + \eta_4\frac{(x-h_t)^4 - 6v^2\tau(x-h_t)^2 + 3v^4\tau^2}{v^8\tau^4}\right\}\phi\left(x; h_t, v^2\tau\right) \end{aligned} \tag{8}$$

Based on the underlying distribution given in Eq. 8, we possibly evaluate the options by

$$w^{CC}(x_t, t) = w^{BS}(x_t, t) + \eta_4 J_4(x_t, t) + \eta_2 J_2(x_t, t), \tag{9}$$

where we have J_4 and J_2 as follows.

$$\begin{aligned} J_4(x_t, t) = &\frac{1}{v^2}\left(r + \frac{1}{2}v^4\right)J_0 + \frac{r}{v^6}\left(4r^2 + v^4\right)Ke^{-r\tau}N(d_2) \\ &+ \frac{K}{4v^5\tau^2\sqrt{\tau}}\left\{\left(12r^2 + v^4\right)\tau^2 - 4\left(v^2 + 2r\ln\frac{x_t}{K}\right)\tau + 4\left(\ln\frac{x_t}{K}\right)^2\right\} \\ &e^{-r\tau}n(d_2) \end{aligned} \tag{10}$$

and

$$J_2(x_t, t) = \frac{1}{v^2}\left(r + \frac{1}{2}v^2\right)J_0 + \frac{2r}{v^2}Ke^{-r\tau}N(d_2) + \frac{K}{v\sqrt{\tau}}e^{-r\tau}n(d_2). \tag{11}$$

CONCLUDING REMARK

The seminal Black Scholes PDE (1) gives the fair price value 1,804 of the call optimal. Each of the three machine learning approaches

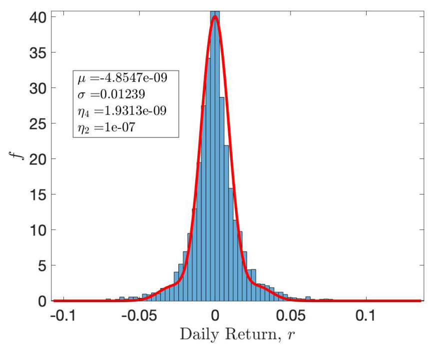

Figure 5: Gram-Charlier Series A type fit for the KOSPI 200 return distribution. The return distribution is shown histogram and its Gram-Charlier fit under Eq. 8 is given red line.

of reinforcement, deep, and EM learnings is performed 100 times. The fair prices at-the-money due to these simulations are 1,820, 1,770, and 1,835, respectively. Our analytic solution in Eq. (9) using the heavy-tailed Gram-Charlier A series distribution as the underlying distribution results in its optimal price 1,818 at-the-money. Among the three machine learning results, the reinforcement learning fits the most to the explainable solution. It allows us to select the reinforcement learning as an optimal one among them.

This makes us to understand better how reinforcement learning presents the optimal price for the heavy-tailed returns of the KOSPI. We believe our contribution is directly related to the comparison between the machine learning prices with the analytical, which shows reinforcement learning generates best result under heavy-tailed distribution. Then, it gives us inductive reasonings to explain the learning.

Particularly it includes risk in options, which is ignored in "risk-neutral" Black-Scholes pricing. Hedging is on the way to the pricing in the reinforcement case, and this effect is captured in our heavy-tailed solution. It recognizes discrepancies among hedging times, which make the return distribution leptokurtic.

REFERENCES

[1] B.S. Choi and M.Y. Choi. 2018. General solution of the Black–Scholes boundary-value problem. *Physica A: Statistical Mechanics and its Applications* 509 (2018), 546 – 550.

[2] C.J. Corrado and T. Su. 1996. Skewness and kurtosis in S&P 500 Index returns implied by option prices. *Financial Research* 19 (1996), 175–192.

[3] C.J. Corrado and T. Su. 1997. Implied volatility skews and stock return skewness and kurtosis implied by stock option prices. *European Journal of Finance* 3 (1997), 73–85.

[4] Eugene F. Fama. 1965. The Behavior of Stock-Market Prices. *The Journal of Business* 38, 1 (1965), 34–105.

[5] I. Halperin. 2017. QLBS: Q-Learner in the Black-Scholes(-Merton) Worlds. *arXiv.org* (2017), 1712.04609.

[6] J. Han, A. Jentzen, and Weinan E. 2018. Solving high-dimensional partial differential equations using deep learning. *Proceedings of the National Academy of Sciences* 115, 34 (2018), 8505–8510. https://doi.org/10.1073/pnas.1718942115

[7] E. Jondeau, S.-H. Poon, and M. Rockinger. 2007. *Financial Modeling Under Non-Gaussian Distributions.* Springer-Verlag, London.

[8] B. Mandelbrot. 1963. The Variation of Certain Speculative Prices. *The Journal of Business* 36, 4 (1963), 394–419.

ACKNOWLEDGMENTS

This work was supported by Korea Institute of Science and Technology's Economic and Social Science Research Initiative 2E29244 and Pride-up project. The authors also declare no conflict of interest.