A Graph Transformer Framework for HK-Mainland Cross-Border Option Pricing with Dynamic Correlation Modeling

Abstract. Introduction of a novel framework that uses Graph-Transformer model on the PDE solving for pricing cross-border derivatives between Hong Kong and Mainland China market. We model the financial ecosystem as a dynamic graph where node represents A-Share and H-share while edges captures the Stock Connect programs linkage with North/South capital flows as edge features.

Keywords: Graph-Transformer method · Option Pricing

1 Introductions

Considering the extended multi-asset black-scholes PDE :

$$\frac{\partial V}{\partial t}+\frac{1}{2}\sigma_A^2 S_A^2\frac{\partial^2 V}{\partial S_A^2}+\frac{1}{2}\sigma_H^2 S_H^2\frac{\partial^2 V}{\partial S_H^2}+\rho_{AH}\sigma_A\sigma_H S_A S_H\frac{\partial^2 V}{\partial S_A\partial S_H}+rS_A\frac{\partial V}{\partial S_A}+rS_H\frac{\partial V}{\partial S_H}-rV=0$$

While in our model, we consider :

$$\rho_{AH}(t)=f_\theta(\mathcal{G}(t))$$

s.t Its dynamic and graph-dependent :
Where :

- $\mathcal{G}(t)=(V,E,X(t),A(t))$ is the financial graph at time t
- f_θ is our Graph Transformer model
- $X(t)$ are node features (prices, volumes, etc.)
- $A(t)$ are edge features (capital flows)

2 Graph-Transformer Architecture

Now considering the architecture of the model :
Initial Encoding :

$$h_i^{(0)}=MLP([S_i,V_i,F_i^{north},F_i^{south},\cdots])$$

Multi-Head Graph Attention :

$$h_i^{(l+1)}=||_{k=1}^{K}\sigma(\sum_{j\in\mathcal{N}(i)}\alpha_{ij}^{k,(l)}W^{k,(l)}h_j^{(l)})$$

Attention Coefficients :

$$a_{ij}^{k,(l)}=\frac{exp(LeakyReLU(a^{k,(l)T}[W^{k,(l)}h_i^{(l)}||W^{k,(l)}h_j^{(l)}||e_{ij}]))}{\sum_{m\in\mathcal{N}(i)}exp(LeakyReLU(a^{k,(l)T}[W^{k,(l)}h_i^{(l)}||W^{k,(l)}h_m^{(l)}||e_{im}]))}$$

Temporal Transformer:
For temporal pattern $\mathbf{H}=[\mathbf{h}^{(t-T)},\ldots,\mathbf{h}^{(t)}]$:

$$Z=Transformer(H+P)$$

where $\mathbf{P}$ is positional encoding.

3 Correlation Predictions

Edge-wise Correlations :
For edge (i,j) connecting A-share i and H-share j:

$$p_{ij}(t)=tanh(w_p^T[z_i^{(t)}||z_j^{(t)}+b_\rho])$$

where $\mathbf{z}_i^{(t)}$ is the final node embedding after spatial-temporal processing.

4 PDE-Constrained Learning

We can formulate this as a Physics-Informed Neural Network (PINN):
Loss Function :

$$\mathcal{L} = \mathcal{L}_{data} + \lambda \mathcal{L}_{PDE}$$

Data Loss :

$$\mathcal{L}_{data} = \frac{1}{N} \sum_{i=1}^{N} (V_{model}(S_A^i, S_H^i, \rho_{graph}^i) - V_{market}^i)^2$$

PDE Loss :

$$\mathcal{L}_{PDE} = \frac{1}{M} \sum_{j=1}^{M} ||BS - PDE(V, S_A^j, S_H^j, \rho_{graph}^j)||^2$$

where the PDE residual is computed using automatic differentiation.

5 Option Pricing with Graph Correlation

Spread Option Price:
Using Kirk's approximation with graph correlations :

$$V_{spread}(S_A, S_H, K, T) = e^{-rT}[F_A N(d_1) - (F_H + K)N(d_2)]$$

Where :

- $F_A = S_A e^{rT}$, $F_H = S_H e^{rT}$
- $\sigma_{\text{spread}} = \sqrt{\sigma_A^2 + \sigma_H^2 - 2\rho_{AH}(t)\sigma_A\sigma_H}$
- $d_1 = \frac{\ln(F_A/(F_H+K)) + \frac{1}{2}\sigma_{\text{spread}}^2 T}{\sigma_{\text{spread}}\sqrt{T}}$
- $d_2 = d_1 - \sigma_{\text{spread}}\sqrt{T}$

6 Stochastic Correlation Process :

Notice that :

$$d\rho_{AH}(t) = \mu_\rho(\mathcal{G}(t))dt + \sigma_\rho(\mathcal{G}(t))dW_\rho$$

Where both drift and volatility are outputs :

$$\mu_\rho, \sigma_\rho = g_\theta(\mathcal{G}(t))$$

Leading to a system of coupled SDEs :

$$dS_A = \mu_A S_A dt + \sigma_A S_A dW_A$$

$$dS_H = \mu_H S_H dt + \sigma_H S_H dW_H$$

$$d\rho_{AH} = \mu_\rho(\mathcal{G})dt + \sigma_\rho(\mathcal{G})dW_\rho$$

$$< dW_A, dW_H >= \rho_{AH} dt$$

7 Results

Considering the results, our model has achived a stable convergence with learning rate scheduling, without any overfitting observed while the error is highly accurate, with tested MAE : 0.010812, MSE : 0.000146.

Considering the economically predictions fall into the range of 0.77 and 0.83 across A-H pairs, the graph-driven correlation produced option value with the range \$0.82 to 2.48 for spread options and \$1.28-4.19 for basket options. Showing that spread option price should increase by 540% as correlation move from 0.1 to 0.9.

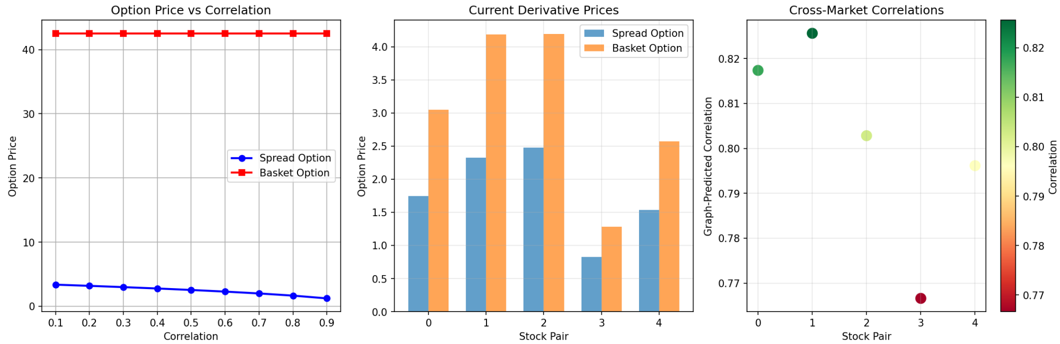

Fig. 1. Economical Results

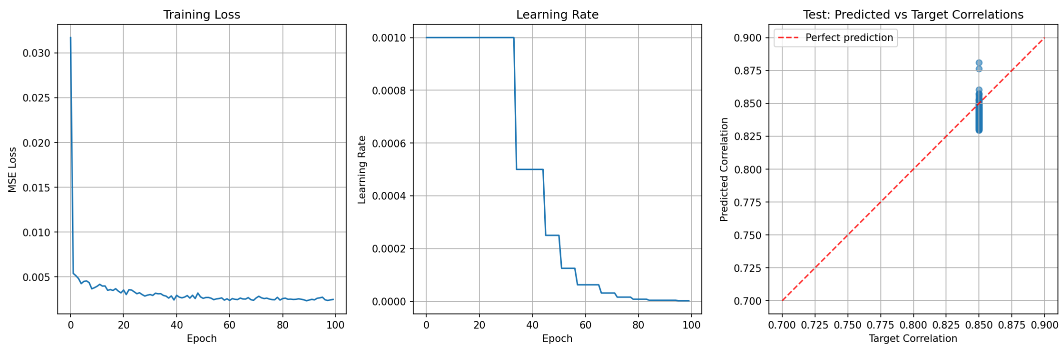

Fig. 2. Training Results

References

1. Lin, C.-I., Chiang, K.-Y., Lee, M.-C., Lee, Y.-H. (2021). The Cross-Border Price Discovery and the Shanghai-Hong Kong Stock Connect. Advances in Management and Applied Economics, 13–28. https://doi.org/10.47260/amae/1132
2. Maeda, Minoru, and Masaharu Kojima. Vecer, J. (2012). BLACK-SCHOLES REPRESENTATION FOR ASIAN OPTIONS. Mathematical Finance, 24(3), 598–626. https://doi.org/10.1111/mafi.12012
3. Velickovic, P., Cucurull, G., Casanova, A., Romero, A., Lio', P., Bengio, Y. (2018). Graph Attention Networks. ICLR. https://doi.org/10.17863/CAM.48429
4. Yun, S., Jeong, M., Kim, R., Kang, J., Kim, H. J. (2020, February 4). Graph Transformer Networks. ArXiv.org. https://doi.org/10.48550/arXiv.1911.06455