Data-Driven Models for Discrete Hedging Problem: From One-Step to Multi-Steps Hedging

Ke Nian
Supervisors: Prof.Yuying Li and Prof.Thomas.F.Coleman

David R. Cheriton School of Computer Science,
University of Waterloo,

Waterloo, Canada

January 26, 2020

Agenda

Black-Scholes Model

- Geometric Brownian Motion:

$$\frac{dS}{S} = \mu dt + \sigma dZ$$

- Price of the option: $V(S,t)$
- From Ito's Lemma:

$$dV = (\mu S \frac{\partial V}{\partial S} + \frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2})dt + \sigma S \frac{\partial V}{\partial S} dZ$$

Black-Sholes Partial Differential Equation

Set up a hedging portfolio

- A short position in an option $-V$
- Long $\frac{\partial V}{\partial S}$ shares of S

$$\Pi = -V + \frac{\partial V}{\partial S} S$$

Thus the random ΔZ will be canceled:

$$\begin{aligned}\Delta \Pi &= -\Delta V + \frac{\partial V}{\partial S}\Delta S \\ &= -(\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2})\Delta t \\ &= r\Pi\Delta t = r(-V + \frac{\partial V}{\partial S} S)\Delta t \text{ (No Arbitrage)}\end{aligned}$$

Black-Scholes Partial Differential Equation:

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

Set Up Self-Financing Portfolio

Consider a portfolio P_t which is composed of:

- A short position on option V_t
- Long α_t (hedging position) shares of S_t
- An amount in a risk-free bank account B_t

The hedging portfolio is rebalanced at discrete times t_i. The hedging position is given by α_{t_i} Initially, we have

$$P_{t_0} = -V_{t_0} + \alpha_{t_0} S_{t_0} + B_{t_0} = 0$$

And

$$B_{t_0} = V_{t_0} - \alpha_{t_0} S_{t_0}$$

Rebalance Self-Financing Portfolio

At each rebalancing time t_i, we update our hedging position by change the share we hold from $\alpha_{t_{i-1}}$ to α_{t_i} at t_i, where any required cash is borrowed, and any excess cash is loaned. Assume $\Delta t = t_i - t_{i-1}$ is fixed. The bank account is updated by:

$$B_{t_i} = e^{r\Delta t} B_{t_{i-1}} - S_{t_i}(\alpha_{t_i} - \alpha_{t_{i-1}})$$

Hedging Objective

Let t_i^+ and t_i^- to be the time immediately after and immediately before t_i. Assume that the performance is measured at the t_N:

$$
\begin{aligned}
P_{t_N^-} &= e^{r\Delta t} B_{t_{N-1}} - V_{t_N} + S_{t_N} \alpha_{t_{N-1}} \\
&= \sum_{j=0}^{N-1} \left\{ \left[e^{r(N-j-1)\Delta t} S_{t_{j+1}} - e^{r(N-j)\Delta t} S_{t_j} \right] \alpha_{t_i} \right\} \\
&+ e^{rN\Delta t} V_{t_0} - V_{t_N}
\end{aligned}
$$

If we always set $\alpha = \frac{\partial V}{\partial S}$ and let $\Delta t \to 0$ (we continuously rebalance the portfolio), then $P_{t_N^-} = 0$. In reality, we can only rebalance discretely and $P_{t_N^-}$ can take positive (profit) and negative value (loss).

Practitioner Black-Scholes (BS) Delta Hedging

- BS model:

$$\frac{dS}{S} = rdt + \sigma dZ$$

$$\sigma: \text{ Constant}$$

- Implied volatility

$$\sigma_{imp} = V_{BS}^{-1}(V_{mkt}, .)$$

V_{mkt}: market option price
V_{BS}^{-1} : inverse of BS pricing function

- Use BS Delta with implied volatility as hedging position:

$$\delta_{BS} = \frac{\partial V_{BS}}{\partial S}$$

Problem with Black-Scholes Delta

Problem with the traditional Black-Scholes delta:

- Market violates Black-Scholes assumption
- Dependence of implied volatility on underlying asset price

Variants of delta hedging strategy:

- Stochastic Volatility Model
- Local Volatility Model
- Minimum Variance Approach
- **Data-Driven Approach**

Minimum Variance Approach

The correction for the dependence of implied volatility on asset price:

- The Minimum Variance (MV) delta:

$$\delta_{MV} = \frac{\partial V_{BS}}{\partial S} + \frac{\partial V_{BS}}{\partial \sigma_{imp}} \frac{\partial \sigma_{imp}}{\partial S}$$

- A parametric model [1]learned from market data can be used to estimate $\frac{\partial \sigma_{imp}}{\partial S}$
- Local volatility model and stochastic volatility model (e.g. SABR) can also be used to calculate the $\frac{\partial \sigma_{imp}}{\partial S}$.

[1]Hull, J. and White, A., "Optimal delta hedging for options." Journal of Banking and Finance 82 (2017): 180-190.

Problem with Parametric Approach

Parametric approaches:

- Model mis-specification.
- Sub-optimal for discrete hedging problems.

Data-driven approaches:

- Minimum assumptions on S.
- Model is determined by market data.

The indirect data-driven approach [2]has been proposed:

- Determine the data-driven pricing function $V(X)$ using regression model.
- Compute $\frac{\partial V(X)}{\partial S}$ as hedging position

[3]Hutchinson, J.M., Lo, A.W. and Poggio, T., "A nonparametric approach to pricing and hedging derivative securities via learning networks." The Journal of Finance 49.3 (1994): 851-889.

Motiviation of Direct Data-Driven Approach

The indirect data-driven approach has the following problems:

- Unnecessary intermediate procedure.
- Sub-optimal for discrete hedging.
- Model parameters depend on the asset price.

Direct data-driven approach can be more useful in practice.

- Customized hedging position function.
- Directly learn the hedging position.

David R. Cheriton School of Computer Science, University of Waterloo

Direct Data-Driven Local Hedging Approach

The direct data-driven approach is

$$\min_{f}\left[\frac{1}{N}\sum_{i=1}^{N}(\Delta V_i - \Delta S_i f(X_i))^2\right]$$

- ΔV_i : the change of option value in data instance i.
- ΔS_i : the change of asset price in data instance i.
- $f(X_i)$: option hedging position function.
- Data-driven models outperform other delta hedging strategies[3].

[4]Nian, Ke, Thomas F. Coleman, and Yuying Li. "Learning minimum variance discrete hedging directly from the market." Quantitative Finance (2018): 1-14.

Understanding the Local Hedging Objective

Recall the hedging portfolio value at t_N is:

$$
\begin{aligned}
P_{t_N^-} &= e^{r\Delta t} B_{t_{N-1}} - V_{t_N} + S_{t_N}\alpha_{t_{N-1}} \\
&= \sum_{j=0}^{N-1} \left\{ \left[e^{r(N-j-1)\Delta t} S_{t_{j+1}} - e^{r(N-j)\Delta t} S_{t_j} \right] \alpha_{t_i} \right\} \\
&+ e^{rN\Delta t} V_{t_0} - V_{t_N}
\end{aligned}
$$

Assume $r = 0$ and we evaluate performance at t_1:

$$
\begin{aligned}
P_{t_1^-} &= (S_{t_1} - S_{t_0})\alpha_{t_0} - (V_{t_1} - V_{t_0}) \\
&= \Delta S \alpha_{t_0} - \Delta V
\end{aligned}
$$

Local hedging objective corresponds to one-step hedging.

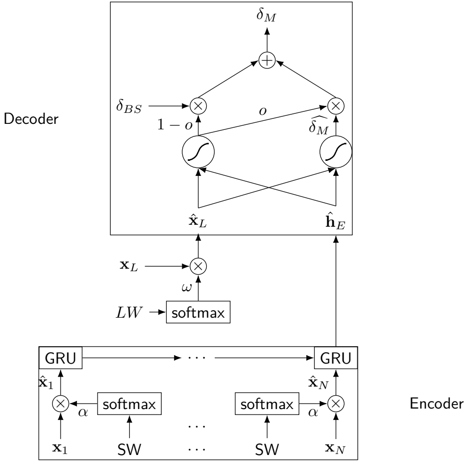
Volatility Clustering and Financial Time Series

Sequential learning framework may further improve the performance:

- Volatility clustering observed in the financial market.
- Autocorrelation between data instances near in time.
- Dependence of option pricing function on the past history of the underlying has been shown in GARCH models [4].

[5] Heston, Steven L., and Saikat Nandi "A closed-form GARCH option valuation model." The review of financial studies 13.3 (2000): 585-625.

Encoder-Decoder Model



Evaluation Criteria: Local Risk

The percentage increase in the effectiveness over the BS hedging:

$$Gain = 1 - \frac{SSE[\Delta V_i - \Delta S_i \delta^i]}{SSE[\Delta V_i - \Delta S_i \delta^i_{BS}]}$$

- SSE: sum of squared errors
- δ: hedging position computed from different models
- δ_{BS}: BS delta

Experimental Setting

- Data: S&P 500 index option from Jan 2007 to Aug 2015
- The models to be compared:
  - $\mathrm{DKL}_{\mathsf{SPL}}$: Direct data-driven kernel learning model.
  - MV: Minimum variance hedging formula.
  - LVF: Local volatility function model.
  - SABR: SABR stochastic volatility model.
  - DRNN: The proposed encoder-decoder model

Call Option Daily Hedging

Delta	MV (%)	SABR(%)	LVF(%)	Data-Driven Model DKL$_{SPL}$ (%) Traded	DKL$_{SPL}$ (%) All	DRNN (%) Traded	DRNN (%) All
0.1	42.1	39.4	42.6	47.1	48.6	32.3	33.8
0.2	35.8	33.4	36.2	37.8	40.0	33.7	36.4
0.3	31.1	29.4	30.3	34.1	35.1	34.1	**35.5**
0.4	28.5	26.3	26.7	32.3	32.0	**33.7**	**34.2**
0.5	27.1	24.9	25.5	29.3	29.4	**35.1**	**33.0**
0.6	25.7	25.2	25.2	29.9	28.4	**35.6**	**32.1**
0.7	25.4	24.7	25.8	29.0	26.8	**31.8**	**29.7**
0.8	24.1	23.5	25.4	25.9	24.7	**28.6**	**26.5**
0.9	16.6	17.0	16.9	17.7	13.9	**19.3**	**18.9**
Overall	25.7	24.6	25.5	31.3	26.0	**32.9**	**28.7**

- Performance will be slighted better than $\mathrm{DKL}_{\mathrm{SPL}}$.

Call Option Weekly Hedging and Monthly Hedging

Delta	Data-Driven Model $\mathrm{DKL_{SPL}}$(%) Traded	All	DRNN(%) Traded	All
0.1	38.9	38.3	**47.8**	**45.6**
0.2	29.0	26.9	**48.5**	**46.0**
0.3	23.5	25.3	**48.5**	**46.6**
0.4	20.8	24.3	**45.9**	**45.4**
0.5	19.9	22.8	**46.6**	**45.0**
0.6	17.3	19.5	**44.8**	**43.1**
0.7	16.8	17.7	**43.9**	**42.4**
0.8	12.5	12.3	**37.7**	**39.0**
0.9	6.2	5.1	**16.4**	**29.1**
Overall	20.2	17.1	**43.7**	**40.5**

Delta	Data-Driven Model $\mathrm{DKL_{SPL}}$ (%) Traded	All	DRNN (%) Traded	All
0.1	22.7	24.8	**53.9**	**39.4**
0.2	23.5	25.5	**51.7**	**48.3**
0.3	24.0	24.6	**50.2**	**49.1**
0.4	21.0	20.7	**47.8**	**48.3**
0.5	13.5	12.7	**44.5**	**47.6**
0.6	14.3	13.5	**44.6**	**47.4**
0.7	6.1	7.0	**35.3**	**42.9**
0.8	5.3	4.1	**24.8**	**34.1**
0.9	4.1	2.3	**10.5**	**19.9**
Overall	16.3	12.5	**44.5**	**42.3**

Table: Weekly(Left) and Monthly(Right)

- Performance will be significantly better than $\mathrm{DKL_{SPL}}$.

From One-Step hedging to Multi-Steps Hedging

In practice, multi-steps hedging is more common. Particularly, the traders usually would like to hedge to the expiry of the option. In other words, $t_N = T$ and V_{t_N}=payoff.

$$
\begin{aligned}
P_{t_N^-} &= e^{r\Delta t} B_{t_{N-1}} - V_{t_N} + S_{t_N} \alpha_{t_{N-1}} \\
&= \sum_{j=0}^{N-1} \left\{ \left[e^{r(N-j-1)\Delta t} S_{t_{j+1}} - e^{r(N-j)\Delta t} S_{t_j} \right] \alpha_{t_i} \right\} \\
&+ e^{rN\Delta t} V_{t_0} - V_{t_N}
\end{aligned}
$$

Total Hedging Objective

Assume we have M samples of sequences. Each sequence is of length N. Let the hedging position given by a function $\delta_t = f(X_t, y_t)$. The objective of optimizing f is:

$$\min_f \quad \frac{1}{2M} \sum_{i=1}^{M} (P^i_{t_N^-})^2$$

Where $P^i_{t_N^-}$ is the portfolio at t_N^- of sample i:

$$P^i_{t_N^-} = \sum_{j=0}^{N-1} \left\{ \left[e^{r(N-j-1)\Delta t} S^i_{t_{j+1}} - e^{r(N-j)\Delta t} S^i_{t_j} \right] f(\mathbf{X}^i_{t_j}, \mathbf{y}^i_{t_j}) \right\} + e^{rN\Delta t} V^i_{t_0} - V^i_{t_N}$$

Creating Training Samples and Testing Samples

The training and testing sample can be generated by simulation:

- Black-Scholes Model:
  - Analytical delta.
  - Analytical optimal total hedging position.[6]
  - Total hedging position based on spline function.[6]
- Heston Model:
  - Analytical delta.

[6] Thomas F Coleman, Yuying Li, and Maria-Cristina Patron. " Total risk minimization using monte carlo simulations". Handbooks in Operations Research and Management Science, 15:593–635, 2007.

Evaluation Criteria

The evaluation criteria are:

- Total risk: Average of the absolute value of the final portfolio value:

$$e^{-rN\Delta t}\frac{\sum_{i=1}^{M}|P^i_{t_N^-}|}{M}$$

- Total cost: Average of the total cost in rebalancing the portfolio:

$$e^{-rN\Delta t}\frac{\sum_{i=1}^{M}(e^{rN\Delta t}V^i_{t_0}-P^i_{t_N^-})}{M}$$

David R. Cheriton School of Computer Science, University of Waterloo

Data-Driven Hedging Model

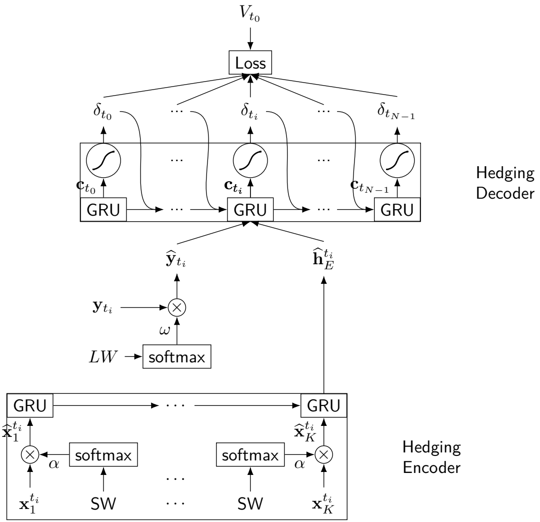


Figure: Model For Synthetic Experiments

David R. Cheriton School of Computer Science, University of Waterloo

Synthetic Case: Black-Scholes Model

We have 600 time steps with $\Delta t = 1/600$. We can hedge every 25,50,100,300 time steps. $S = K = 100$

method	bi-weekly	monthly	quarterly	semi-annually
GRU	0.8376	1.1426	1.6896	2.8038
Spline	0.8563	1.1789	1.6518	2.7843
Analytical	0.8295	1.1636	1.6479	2.7914
BS	0.9481	1.3385	1.9128	3.4582

Table: Total Risk

method	bi-weekly	monthly	quarterly	semi-annually
GRU	5.9682	5.8370	5.8549	5.1759
Spline	5.9118	5.8445	5.7119	5.2530
Analytical	5.9413	5.8773	5.7399	5.2565
BS	6.0483	6.0897	6.1734	6.5382

Table: Total Cost

Synthetic Case: Heston Model

The parameter for heston model: $r = 0.02$, $\overline{v} = 0.04$, $\kappa = 1.15$, $\eta = 0.39$, $S = 100$, $K = 100$, $v = 0.04$, $\rho = -0.64$, $\tau = 1$.

method	bi-weekly	monthly	quarterly	semi-annually
GRU	1.9907	2.2183	2.5345	3.5552
Heston	2.6228	2.8492	3.2501	4.1049

Table: Total Risk (Heston Model)

method	bi-weekly	monthly	quarterly	semi-annually
GRU	8.3632	8.3601	8.3387	8.3941
Heston	8.3770	8.3676	8.3894	8.3897

Table: Total Cost (Heston Model)

Challenges of Obtaining Real Data

1. Only one single path for underlying asset
   - Extract segments from the path.
2. Options with specific K and expiry are not traded every day.
   - Use a calibrated price surface to fill the missing data.
3. Options in real market only have fixed expiry dates:
   - Use a calibrated price for option with expiries not seen on market.

David R. Cheriton School of Computer Science, University of Waterloo

Call-Put Parity

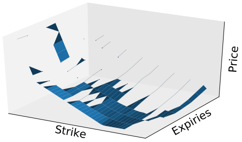
We can also use the Call-Put parity to increase the number price observed from market:

$$C - P = S - DK$$

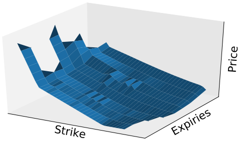
where C is the (current) value of a call, P is the (current) value of a put, D is the discount factor, and K is the strike price.

Volatility Interpolation Illustration

The purpose of volatility Interpolation[5] is to create a price surface that can be used to obtain option price unobserved from market:



(a) Before



(b) After

Figure: Illustration of Constructing a Price Surface

[7] Jesper Andreasen and Brian Huge. "Volatility interpolation." Risk, 24(3):76, 2011.

Dupire's forward equation

- The forward price of a call option for delivery at time T: $C(T,K)$
- The spot price at t is: $C(T,K)e^{-\int_t^T r(s)ds}$.
- It can be shown that :

$$\frac{\partial C(T,K)}{\partial T} = \frac{1}{2}\sigma^2(T,K)K^2\frac{\partial^2 C(T,K)}{\partial K^2}$$

Model Calibration

We can write finite difference discretization of the Dupure forward equation as:

$$M \begin{bmatrix} C(T_i, K_0) \\ C(T_i, K_1) \\ C(T_i, K_2) \\ \vdots \\ C(T_i, K_{n-1}) \\ C(T_i, K_n) \end{bmatrix} = \begin{bmatrix} C(T_{i+1}, K_0) \\ C(T_{i+1}, K_1) \\ C(T_{i+1}, K_2) \\ \vdots \\ C(T_{i+1}, K_{n-1}) \\ C(T_{i+1}, K_n) \end{bmatrix}$$

We try to find M that so that $C(T_{i+1}, K_j) = C_{mkt}(T_{i+1}, K_j)$. This can be done by:

$$\inf_{\sigma(T_i,.)} \sum_j \left(\frac{C(T_{i+1}, K_j) - C_{mkt}(T_{i+1}, K_j)}{Vega_{bs}^{mkt}(T_{i+1}, K_j)} \right)^2$$

Note that, for each T_i, we solve a separate optimization.

Interpolation Over the Domain of Expiries

After optimization, the local volatility functions are translated into arbitrage-consistent prices for a discrete set of expiries but it does not directly specify the option prices between the expiries. We can fill in the gaps by:

$$\frac{C(T,K)-C(T_i,K)}{T-T_i}=\frac{1}{2}\sigma(T_i,K)^2K^2\frac{\partial^2 C(T_{i+1},K)}{\partial K^2}, T\in[T_i,T_{i+1})$$

Benefits of Interpolation Based On Local Vol Model

Interpolation based on the above procedure can guarantee the option price given by interpolation is arbitrage-free:

1. No call spread arbitrage:

$$\frac{\partial C(T,K)}{\partial K} \leq 0$$

2. No butterfly spread arbitrage:

$$\frac{\partial^2 C(T,K)}{\partial K^2} \geq 0$$

3. No calendar spread arbitrage:

$$\frac{\partial C(T,K)}{\partial T} \geq 0$$

David R. Cheriton School of Computer Science, University of Waterloo

Total Hedging Model for Real Data Case

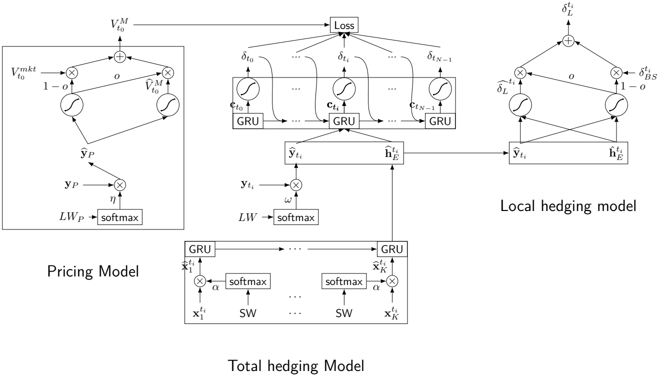


Figure: Refined Model For Real Cases

Primitive Experimental Setting

- Testing period is from 2007 and 2014 for SP500 index option.
- Scenario: weekly hedging for two months.
- All data in previous years is used as training.
- Model are updated yearly.
- Early stopping is used as regulation
- Performance is evaluated with **relative** hedging error.

$$rel_{err} = \frac{P^i_{t_N^-}}{V_{t_0}}$$

Value-At-Risk of **Relative** Hedging Error

method	Total	Local	BS
2007	**-0.8622**	-1.044	-2.3724
2008	**-1.1430**	-1.0782	-4.9241
2009	**-0.4563**	-1.3607	-2.3771
2010	**-0.4509**	-0.6817	-1.7911
2011	**-0.7062**	-0.9049	-1.9094
2012	**-0.3866**	-1.7635	-2.6473
2013	**-0.4635**	-2.7910	-4.2887
2014	**-1.5424**	-2.0567	-3.1884

Table: Value-At-Risk

Expected Shortfall of **Relative** Hedging Error

method	Total	Local	BS
2007	**-1.1568**	-1.8545	-5.4942
2008	**-2.0683**	-4.9241	-7.3248
2009	**-0.6443**	-2.3772	-5.0323
2010	**-0.6207**	-1.1806	-3.6964
2011	**-1.1439**	-1.9460	-3.2358
2012	**-0.5497**	-3.2662	-4.7711
2013	**-0.6460**	-4.3091	-6.6587
2014	**-1.9005**	-3.4671	-5.2354

Table: Expected Shortfall

Mean Absolute **Relative** Hedging Error

method	Total	Local	BS
2007	**0.3769**	0.7357	1.3396
2008	**0.5034**	0.6852	0.9068
2009	**0.3041**	0.6597	0.5092
2010	**0.3412**	0.5837	1.1331
2011	**0.3507**	0.4611	0.8513
2012	**0.2726**	0.5858	0.8084
2013	**0.3055**	0.8961	0.9710
2014	**0.5876**	0.9509	1.6091

Table: Mean Absolute **Relative** Hedging Error

Standard Deviation of **Relative** Hedging Error

method	Total	Local	BS
2007	**0.4977**	2.2085	3.2548
2008	**0.7645**	2.4953	3.7674
2009	**0.3785**	1.3938	2.6829
2010	**0.4770**	1.1388	1.3975
2011	**0.4787**	0.7269	1.1465
2012	**0.3339**	0.9633	1.5448
2013	**0.3717**	1.4185	2.6968
2014	**0.8899**	2.2879	5.3217

Table: Standard Deviation

Q & A

Thank you very much!

Any Questions?