Learning Minimum Variance Discrete Hedging Directly From Market

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

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

Waterloo, Canada

May 1, 2019

Agenda

Practitioner Black-Scholes (BS) Delta Hedging

- BS model:

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

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

- Implied Volatility

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

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

- BS Delta:

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

Problem with Black-Scholes Delta

Problem with the traditional Black-Scholes Delta:

- Market violates BS assumption
- Dependence of volatility on underlying asset price

Variants of Hedging Strategy:

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

Stochastic Volatility Model

Stochastic volatility models:

- Heston Model
- SABR Volatility Model
- GARCH Model

Minimum Variance Approach

Considering the the dependence of imply volatility on asset price:

- The Minimum Variance (MV) delta:

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

- The authors [1] propose:

$$\frac{\partial E(\sigma_{imp})}{\partial S} = \frac{a + b\delta_{BS} + c\delta_{BS}^2}{S\sqrt{T}} \tag{1}$$

a, b and c are the parameter to be fitted using market data.

[1] Hull,J and White,A ,'Optimal Delta Hedging for Options', unpublished manuscript

Local Volatility Model

The local volatility function (LVF) [2]: volatility is a deterministic function of S and t.

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

Local volatility model can also be used to calculate the $\frac{\partial E(\sigma_{imp})}{\partial S}$

[2] Coleman, T, Y. Kim, Y. Li and A. Verma, 'Dynamic hedging with a deterministic local volatility function model,' Journal of risk, 4 ,1 (2001):63-89

Problem with Parametric Approach

Problem with the previous parametric approaches:

- Assumptions do not hold in market

Data-driven approach can be more useful in practice.

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

Indirect Data-driven Approach

The indirect data-driven approach [3] can be summarized as following:

- Let X the features from market.
  - Asset price S
  - Strike Price K
  - Time to expiration $T - t$
- Determine the data driven pricing function $V(X)$ using regression model.
- Compute

$$\delta_{ID} = \frac{\partial V(X)}{\partial S}$$

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

Problem with Indirect Data-Driven Approach

Problem with the Indirect Data-Driven Approach:

- Intermediate procedure is not necessary.
- Not suitable for weekly and monthly hedging

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

- Customized hedging position function
- Directly compute the hedging position

Direct Data-driven 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
V_i : the option price in data instance i

Regularized Network

- Indirect data-driven approach:

$$\min_{f \in RKHS} \left[\frac{1}{N} \sum_{i=1}^{N} (V_i - f(x_i))^2 + \lambda \|f\|_K^2 \right]$$

- Direct data-driven approach:

$$\min_{f \in RKHS} \left[\frac{1}{N} \sum_{i=1}^{N} (\Delta V_i - \Delta S_i f(x_i))^2 + \lambda \|f\|_K^2 \right]$$

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

Regularized Network (2)

Given

$$f(x) = \sum_{i=1}^{N} \alpha_i K(x, x_i)$$

Indirect data-driven approach:

$$\min_{\alpha} (K\alpha - V)^T (K\alpha - V) + \lambda \alpha^T K \alpha$$

Direct data-driven approach:

$$\min_{\alpha} (DK\alpha - \Delta V)^T (DK\alpha - \Delta V) + \lambda \alpha^T K \alpha$$

Where D is the diagonal matrix with ΔS on its diagonal

Regularized Network (3)

Indirect data-driven approach:

$$\alpha = (K + \lambda_P I)^{-1} V$$

Let $\widetilde{K} = DK$, direct data-driven approach:

$$\widetilde{K}^T(\widetilde{K}\alpha - \Delta V) + \lambda K \alpha = 0$$
$$(\widetilde{K}^T \widetilde{K} + \lambda K)\alpha = \widetilde{K}^T \Delta V$$
$$\alpha = (\widetilde{K}^T \widetilde{K} + \lambda K)^{-1} \widetilde{K}^T \Delta V$$

Cross-Validation

For the indirect data-driven approach :

$$\alpha = (K + \lambda_P I)^{-1} V$$

We can calculate the eigen-decomposition of $K = Q \Lambda Q^T$ and then

$$\alpha = Q(\Lambda + \lambda I)^{-1} Q^T V$$

Given the eigen-decomposition, finding $\alpha(\lambda)$ can be done($O(N^2)$).

Leave-One-Out Cross-Validation

- For each data point x_i, building a model using the remaining $N-1$ data points, and measuring the error for x_i
- It can be further shown that, given the eigen-decomposition of $K = Q\Lambda Q^T$, we can estimate the leave-one-out cross-validation errors in $O(N^2)$.

Modification of the direct data driven model

- For our direct data-driven approach:

$$\alpha = (\widetilde{K}^T \widetilde{K} + \lambda K)^{-1} \widetilde{K}^T \Delta V$$

- If we changed the penalty term from $\alpha^T K \alpha$ to $\alpha^T \alpha$:

$$\alpha = (\widetilde{K}^T \widetilde{K} + \lambda I)^{-1} \widetilde{K}^T \Delta v$$

Modification of the direct data driven model (2)

Let the SVD of $\widetilde{K} = U\Sigma V^T$, the α can be determined by

$$
\begin{aligned}
&(\widetilde{K}^T\widetilde{K} + \lambda I)\alpha = \widetilde{K}^T \Delta V \\
&(V\Sigma^T U^T U\Sigma V^T + \lambda I)\alpha = V\Sigma^T U^T \Delta V \\
&V(\Sigma^T\Sigma + \lambda I)V^T\alpha = V\Sigma^T U^T \Delta V \\
&(\Sigma^T\Sigma + \lambda I)V^T\alpha = \Sigma^T U^T \Delta V \\
&\alpha = V(\Sigma^T\Sigma + \lambda I)^{-1}\Sigma^T U^T \Delta V
\end{aligned}
$$

$(\Sigma^T\Sigma + \lambda I)$ is again a diagonal matrix. Then we can still estimate the leave-one-out cross-validation errors in $O(N^2)$

Kernels

In the following experiments, we use:

- The Gaussian kernel

$$K(x, y) = e^{-\frac{\|x-y\|_2^2}{2\sigma_b^2}}$$

- The Spline Kernel:

$$K(x, y) = \int_{lb}^{+\infty} (x-t)_+^d (y-t)_+^d \, dt + \sum_{k=0}^{d} x^k y^k$$

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

Synthetic Data: Experimental Setting

- Dynamic for asset: Heston Model
- Training data: a 2 year stock path
- Testing Data: 100 independent 6-months stock path
- Methods considered:
  - δ^{BS}: implied volatility BS delta
  - $\textsc{Heston}$: analytic Heston delta
  - $\mathrm{DKL}_{\mathsf{SPL}}$: direct learning a spline kernel hedging function
  - $\mathrm{DKL}_{\mathsf{RBF}}$: directly learning a RBF kernel hedging function
  - $\mathrm{IKL}_{\mathsf{SPL}}$: Indirectly computing $\frac{\partial f}{\partial S}$ using a spline kernel
  - $\mathrm{IKL}_{\mathsf{RBF}}$: Indirectly computing $\frac{\partial f}{\partial S}$ using a RBF kernel

Daily Hedging Performance

Method	Gain (%)	$\mathbf{E}(\lvert\Delta \breve{V} - \Delta S f(\mathbf{x})\rvert)$	Std	VaR	CVaR
δ^{BS}	0.0	0.185	0.286	0.380	0.574
IKL_{RBF}	-3.3	0.171	0.291	0.356	0.566
IKL_{SPL}	-183.3	0.291	0.482	0.669	1.105
DKL_{RBF}	63.1	0.120	0.174	0.251	0.352
DKL_{SPL}	**64.9**	**0.121**	**0.170**	**0.255**	**0.345**
HESTON	63.6	0.121	0.173	0.266	0.360

Table: Daily Hedging Comparison

[1] FS #1: $X = \{\text{MONEYNESS}, \text{ TIME-TO-EXPIRY}\}$
[2] Bold entry indicating best Gain

Real Data Hedging Experiments

- Data: S&P 500 index option from Jan 2007 and Aug 2015
- Model Calibration:
  - SABR: daily calibration
  - LVF: $\frac{\partial E(\sigma_{imp})}{\partial S}$ from implied volatility surface
  - MV: Use a 36 months time window to train
  - $\mathrm{DKL}_{\mathsf{SPL}}$: Use a 36 months time window to train. Models are separately calibrated for different Black-Sholes delta range.

S&P 500 Call Options

Delta	SABR (%)	LVF (%)	MV (%)	$\mathrm{DKL}_{\mathsf{SPL}}$ (%) Leave-One-Out[1] Traded	$\mathrm{DKL}_{\mathsf{SPL}}$ (%) Leave-One-Out[1] All
0.1	42.1	39.4	42.6	**44.1**	**44.4**
0.2	35.8	33.4	36.2	**37.8**	**38.1**
0.3	31.1	29.4	30.3	**33.1**	**33.6**
0.4	28.5	26.3	26.7	**30.9**	**31.3**
0.5	27.1	24.9	25.5	**30.0**	**30.4**
0.6	25.7	25.2	25.2	**29.3**	**29.8**
0.7	25.4	24.7	25.8	**28.4**	**30.2**
0.8	24.1	23.5	25.4	22.5	**28.0**
0.9	16.6	**17.0**	16.9	8.1	12.7
Overall	25.7	24.6	25.5	**31.3**	**26.8**

Table: S&P 500 Call Option Daily Hedging: bold entry indicating best Gain

[1] For each month, the penalties for models are determined by leave-one-out cross validation.

Conclusion

- loosing assumption on the market dynamic is a good practise.
- Data driven approach can leads to better performance.

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

Q & A

Thank you very much!

Any Questions?