Dynamic Hedging Using the LSTM Neural Network

Xin Jing
University of Toronto
xin.jing@mail.utoronto.ca

CSLT, Tsinghua University
August, 2017

Introduction

- *For decades, companies have been employing dynamic hedging with futures to minimize price risks from currencies, commodities, interest rates, etc. Usually, this is done by first using GARCH models to predict the price volatility of the next period, and then calculating the optimal hedge ratio.*
- *However, GARCH models suffer from linearity and parsimony. They are useful for approximating future volatility, but lack accuracy. Moreover, data may not satisfy the requirements of GARCH models.*
- *We used an LSTM neural network to predict the optimal hedge ratio* ***directly*** *from past price changes, skipping the step of volatility prediction.*
- *The LSTM neural network achieved much better performance than the five baseline models in all five markets we tested (EUR, ASD, BP, CAD, S&P 500).*

TASK DESCRIPTION

- *Suppose we have a fixed long position of one unit in the spot market.*
- *In the next period, let the spot price change be* $s^{(t)}$*, the future price change be* $f^{(t)}$*, and the risk-minimizing hedge ratio* $b^{(t)}$ *be the short position we want to take in the futures market in an attempt to offset the change in spot price.*
- *The random return to this portfolio,* $x^{(t)}$*, is* $x^{(t)} = s^{(t)} - b^{(t)} \cdot f^{(t)}$.
- *Our goal is to minimize* $Var(x^{(t)})$ *by predicting the optimal* $b^{(t)}$ *periodically and adjusting our position accordingly.*

Baseline Models

- ***BASELINE A.*** *No hedging (calculate $Var(s^{(t)})$ directly)*
- ***BASELINE B.*** *Naive hedging: take a fixed position of one unit in the spot market and one unit in the futures market (i.e. b = 1)*
- ***BASELINE C.*** *Conventional hedging: use a fixed hedge ratio calculated by $b = \frac{Cov(s,f)}{Var(f)}$, where b is time invariant*
- ***BASELINE D.*** *Dynamic hedging using CCC-GARCH with error-correction terms, where $b^{(t)}$ is adjusted periodically*

$$s^{(t)} = \alpha_s + \beta_s(S_{t-1} - \delta F_{t-1}) + \epsilon_{st}, \quad f^{(t)} = \alpha_f + \beta_f(S_{t-1} - \delta F_{t-1}) + \epsilon_{ft},$$

$$\begin{bmatrix} \epsilon_{st} \\ \epsilon_{ft} \end{bmatrix} \Bigg| \Psi_{t-1} \sim N(0, H_t), \quad H_t = \begin{bmatrix} \sigma^2_{s,t} & \sigma_{sf,t} \\ \sigma_{sf,t} & \sigma^2_{f,t} \end{bmatrix} = \begin{bmatrix} \sigma_{s,t} & 0 \\ 0 & \sigma_{f,t} \end{bmatrix} \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix} \begin{bmatrix} \sigma_{s,t} & 0 \\ 0 & \sigma_{f,t} \end{bmatrix},$$

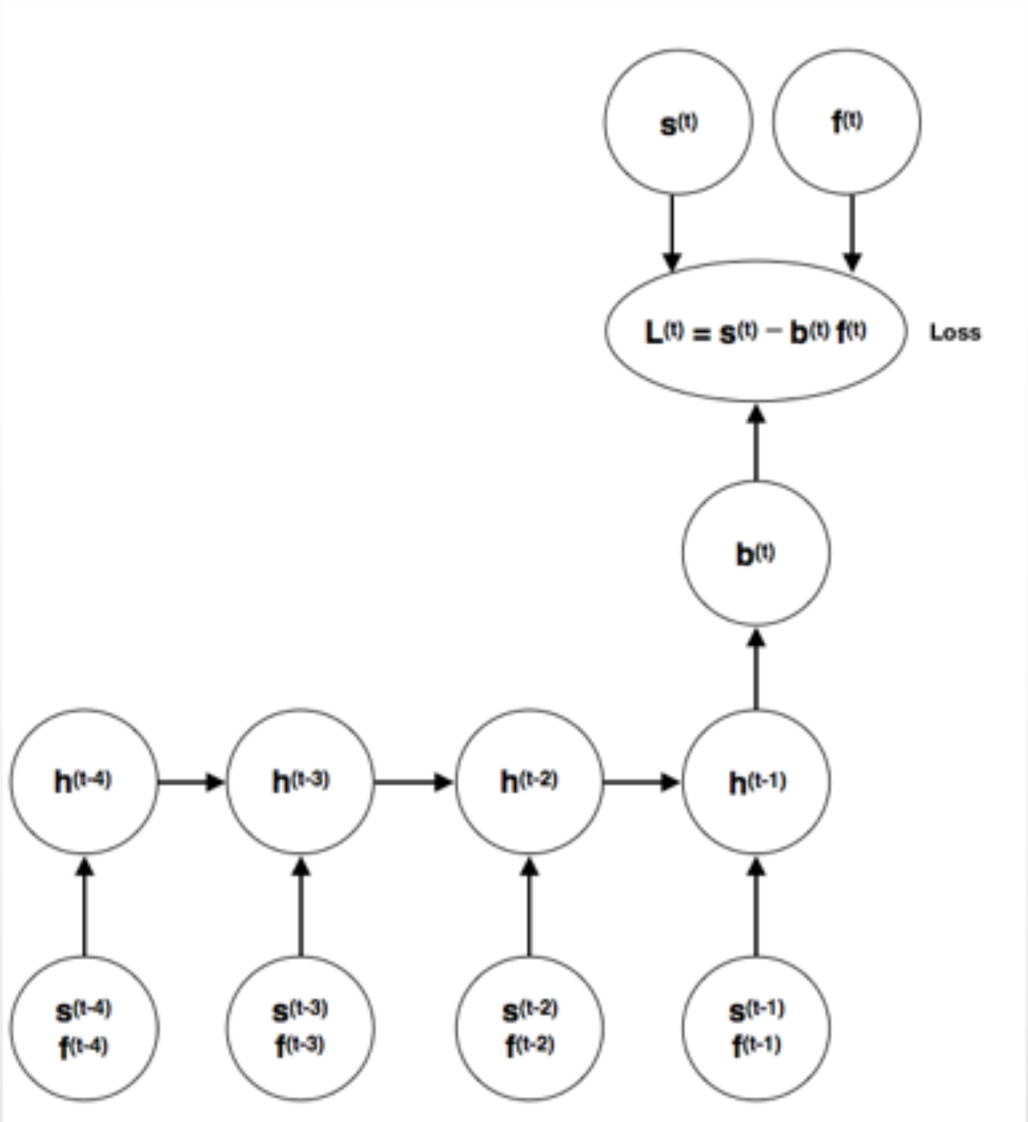
$$\sigma^2_{s,t} = b_{0s} + b_{1s}\epsilon^2_{s,t-1} + b_{2s}\sigma^2_{s,t-1}, \quad \sigma^2_{f,t} = b_{0f} + b_{1f}\epsilon^2_{f,t-1} + b_{2f}\sigma^2_{f,t-1},$$

$$\Rightarrow b^{(t)} = \frac{\sigma_{sf,t}}{\sigma^2_{f,t}},$$

where $S_{t-1}, F_{t-1}, \Psi_{t-1}$ denote the spot price, future price, and the information set at time t-1, respectively.

- ***BASELINE E.*** *Dynamic hedging using DCC-GARCH with error-correction terms, which assumes dynamic conditional correlation between $s^{(t)}$ and $f^{(t)}$; $b^{(t)}$ is adjusted periodically*

Our LSTM Neural Network



- *Nonlinear activation functions inside*
- *Rich in parameters*
- *No requirement on data*
- *Good "memory" of the past information*

$\Big\}$ *Stronger ability in feature extraction and modeling*

Performance

- *The LSTM outperformed all of the five baseline models by achieving smaller variances of portfolio returns.*

Comparison of Hedging Effectiveness on Testing Set

Portforlio Variances	EUR	ASD	BP	CAD	S&P500
No Hedging	2.20E-05	2.63E-05	2.85E-05	1.63E-05	6.54E-05
Naïve Hedge	2.71E-07	7.07E-07	2.47E-07	2.62E-07	6.64E-06
Conventional Hedge	4.36E-07	9.33E-07	2.72E-07	2.41E-07	4.93E-06
Dynamic Hedge with CCC GARCH	3.06E-07	8.03E-07	2.77E-07	2.39E-07	5.65E-06
Dynamic Hedge with DCC GARCH	3.15E-07	8.82E-07	3.17E-07	2.84E-07	5.10E-06
Dynamic Hedge with LSTM	2.64E-07	7.03E-07	2.42E-07	2.29E-07	4.88E-06
Number of hidden states in the LSTM	10	10	15	5	10

Variance Improvement of LSTM hedge compared to :

	EUR	ASD	BP	CAD	S&P500
No Hedging	2.17E-05	2.56E-05	2.83E-05	1.61E-05	6.05E-05
Naïve Hedge	6.77E-09	3.64E-09	5.00E-09	3.30E-08	1.76E-06
Conventional Hedge	1.72E-07	2.30E-07	3.03E-08	1.17E-08	5.29E-08
Dynamic Hedge with CCC GARCH	4.21E-08	1.00E-07	3.54E-08	9.73E-09	7.70E-07
Dynamic Hedge with DCC GARCH	5.06E-08	1.79E-07	7.57E-08	5.41E-08	2.23E-07

- *Daily closing prices of the past 7 years (around 1600 days, with the last 80 days as the testing set)*
- *The LSTM was implemented in Python with TensorFlow and the baseline models were in R.*

References

- Kenneth F. Kroner, Jahangir Sultan. *Time-varying Distributions And Dynamic Hedging with Foreign Currency Futures.* The Journal of Financial and Quantitative Analysis, Vol. 28, No. 4 (Dec., 1993)
- Asim Ghosh. *Hedging with Stock Index Futures: Estimation and Forecasting with Error Correction Model.* The Journal of Futures Markets, Vol. 13, No. 7, 743-752 (1993)
- Donald Lien , Y. K. Tse & Albert K. C. Tsui. *Evaluating the Hedging Performance of the Constant-Correlation GARCH Model.* Applied Financial Economics. 12(11): 791–798. (2002)
- Joey Wenling Yang, David Edmund Allen. *Multivariate GARCH hedge ratios and hedging effectiveness in Australian futures markets.* Accounting and Finance 45 (2004) 301–321
- Luc Bauwens, Sébastien Laurent, Jeroen V. K. Rombouts. *Multivariate GARCH Models: A Survey.* Journal of Applied Econometrics 21: 79–109 (2006)
- Sung Yong Park, Sang Young Jei. *Estimation And Hedging Effectiveness Of Time-Varying Hedge Ratio: Flexible Bivariate GARCH Approaches*. The Journal of Futures Markets, Vol. 30, No. 1, 71–99 (2010)