Deep Hedging

Bent Müller

University of Hamburg, Department of Mathematics

12.6.2023

[1]

[1] Original Paper: [Bühler et al., 2018, Deep Hedging]

Table of Contents

Notation and Setup

- $\Omega := \{\omega_1, \omega_2, \ldots, \omega_N\}$ is our **discrete** set of outcomes.
- $\mathcal{F} := 2^\Omega$ the σ-algebra of all subsets of Ω, so that is $(\Omega, \mathcal{F}, \mathbb{P}, \mathbb{F}, S)$ our financial market.
- $\mathcal{X} := \{X : \Omega \to \mathbb{R}\}$ is the set of all real-valued random variables.
- $\rho : \mathcal{X} \to \mathbb{R}$ is a risk measure.

Notation and Setup

- $l : \mathbb{R} \to \mathbb{R}$ **continuous, convex and non-decreasing** is called a loss function.
- $\rho_l : \mathcal{X} \to \mathbb{R}$, where $\rho_l(X) := \inf_{w\in\mathbb{R}}\{w + \mathbb{E}[l(-X - w)]\}$ defines a convex risk measure, a so-called **Optimized Certainty Equivalent** (OCE) risk measure (Lemma 3.16).
- $S_t^i : \Omega \to \mathbb{R}_{\geq 0}$ is the $\mathcal{F}_t$-measurable price of the i-th risky asset at time t for $0 \leq t \leq T$ and $0 \leq i \leq d$.

Notation and Setup

- $\mathcal{H}^u := \{\phi : \Omega \to \mathbb{R}^{d+1} \mid \forall 0 \leq t \leq T :$ ϕ_{t+1} is $\mathcal{F}_t$-measurable and $\phi_{-1} = \phi_T = 0\}$ is the set of unconstrained hedging strategies.
- $\mathcal{H} \subset \mathcal{H}^u$ is the set of **admissible** hedging strategies which are constrained by e.g. market hours or emittance of certain hedging instruments.
- For simplicity, let $r = 0$.

Trading with Transaction Costs

Let $\delta \in \mathcal{H}$ be a hedging strategy, then we define

$$C_T(\delta) := \sum_{t=0}^{T} c_t(\delta_t - \delta_{t-1})$$

as the **cumulative transaction cost** of trading using δ up to time T. $c_t : \mathbb{R}^{d+1} \to \mathbb{R}^+$ is a non-negative **adapted cost process** satisfying $\forall t : c_t(0) = 0$.

This makes different transaction costs possible, like:

- Proportional Transaction Costs:

  $$c_t(n) = \sum_{i=0}^{d} c_t^i S_t^i |n^i|$$

  for a transaction $n \in \mathbb{R}^{d+1}$.

Trading with Transaction Costs

Let $\delta \in \mathcal{H}$ be a hedging strategy, then we define

$$C_T(\delta) := \sum_{t=0}^{T} c_t(\delta_t - \delta_{t-1})$$

as the **cumulative transaction cost** of trading using δ up to time T. $c_t : \mathbb{R}^{d+1} \to \mathbb{R}^+$ is a non-negative **adapted cost process** satisfying $\forall t : c_t(0) = 0$.

This makes different transaction costs possible, like:

- Fixed Transaction Costs:

$$c_t(n) = \sum_{i=0}^{d} c_t^i 1_{|n^i|>0}$$

for a transaction $n \in \mathbb{R}^{d+1}$.

Trading with Transaction Costs

Let $\delta \in \mathcal{H}$ be a hedging strategy, then we define

$$C_T(\delta) := \sum_{t=0}^{T} c_t(\delta_t - \delta_{t-1})$$

as the **cumulative transaction cost** of trading using δ up to time T. $c_t : \mathbb{R}^{d+1} \to \mathbb{R}^+$ is a non-negative **adapted cost process** satisfying $\forall t : c_t(0) = 0$.

This makes different transaction costs possible, like:

- Or even volatility dependent transaction costs which we won't consider here. In this case c_t might increase with high volatility.

Trading with Transaction Costs

Let $\delta \in \mathcal{H}$ be a hedging strategy, then we define

$$C_T(\delta) := \sum_{t=0}^{T} c_t(\delta_t - \delta_{t-1})$$

as the **cumulative transaction cost** of trading using δ up to time T. $c_t : \mathbb{R}^{d+1} \to \mathbb{R}^+$ is a non-negative **adapted cost process** satisfying $\forall t : c_t(0) = 0$.

This makes different transaction costs possible, like:

- If the agent is allowed to trade in other options to hedge $-Z$, there might be additional costs for high volatility. This is called the **cost of volatility** and can be modeled if we allow c_t to depend on the Black & Scholes **Vega** of each traded asset.

Fundamental Problem

A hedging agent then wishes to achieve the optimal hedge for a given liability $-Z$.

$$\pi(-Z) := \inf_{\delta \in \mathcal{H}} \rho(-Z + (\delta \cdot S)_T - C_T(\delta))$$

$\pi(-Z)$ is the risk of the **optimal hedge** possible when hedging using strategies in $\mathcal{H}$. Then

$$p(Z) = \pi(-Z) - \pi(0)$$

is the **indifference price** at which the agent would be indifferent between holding $-Z$ and not holding any position.

Fundamental Problem

At time T, the agent's terminal wealth is given by

$$p_0 - Z + (\delta \cdot S)_T - C_T(\delta)$$

where p_0 is the amount she received for selling Z in the first place.

Note that p_0 may also be given externally and must not be the indifference price $p(Z)$.

Conditional Value at Risk

Definition (CVaR and VaR)

For any $X \in \mathcal{L}^1(\Omega, \mathcal{F}, \mathbb{P})$ and $\alpha \in (0,1)$, the **Conditional Value at Risk** (CVaR) is defined as

$$\mathrm{CVaR}_\alpha(X) := \frac{1}{\alpha}\int_0^\alpha VaR_u(X)du$$

where $VaR_\lambda(X)$ is the **Value at Risk** at level λ defined as

$$VaR_\lambda(X) := -\inf\{x \in \mathbb{R} \mid \mathbb{P}(X \leq x) \geq \lambda\}.$$

Conditional Value at Risk of S&P 500 Daily Returns

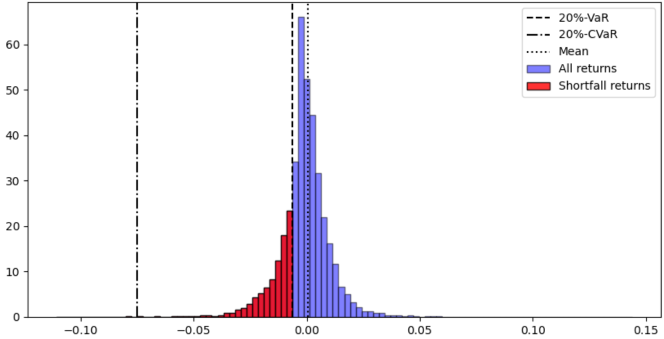

Figure 1: Conditional Value at Risk of S&P 500 Daily Returns since 1993-02-01; mirrored on the 0 for representability

Feed Forward Networks

For $d_i, d_h, d_o \in \mathbb{N}$ and activation function σ, we define a single hidden layer **feed forward network** (FFN) as

$$\mathrm{FFN}^{d_i,d_h,d_o}_{W_2,b_2,W_1,b_1} : \mathbb{R}^{d_i} \to \mathbb{R}^{d_o},\ x \mapsto W_2 \cdot \sigma(W_1 \cdot x + b_1) + b_2.$$

Where

- d_i, d_h, d_o are the dimensions of the input, hidden and output layer respectively,
- $W_1 \in \mathbb{R}^{d_h \times d_i}, W_2 \in \mathbb{R}^{d_o \times d_h}$ are weight matrices and
- $b_1 \in \mathbb{R}^{d_h}, b_2 \in \mathbb{R}^{d_o}$ are bias vectors.

We shorten the parameters and write $\mathrm{FFN}^{d_i,d_h,d_o}_{\theta}$.

Common Activation Functions

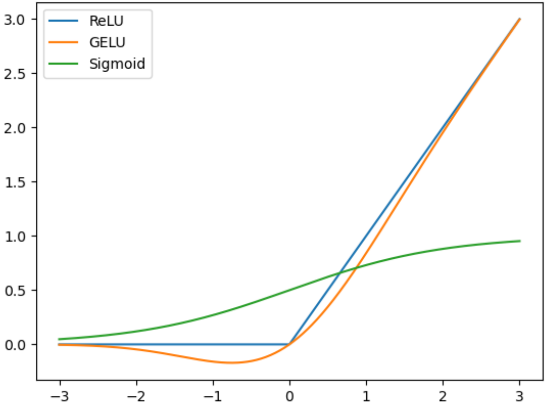

Figure 2: Common Activation Functions

Universal Approximation Theorem

Let σ be non-constant and bounded, $d \in \mathbb{N}$ and $\mathcal{NN}_\infty := \{\mathrm{FFN}_\theta^{d,L,1} \mid L \in \mathbb{N}\}$ be the set of FFNs with arbitrary hidden layer size. Then we have:

1. For any finite measure μ on $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ and $p \in [1, \infty)$, the set $\mathcal{NN}_\infty$ is dense in $L^p(\mathbb{R}^d, \mu)$.

Universal Approximation Theorem

Let σ be non-constant and bounded, $d \in \mathbb{N}$ and $\mathcal{NN}_\infty := \{\mathrm{FFN}_\theta^{d,L,1} \mid L \in \mathbb{N}\}$ be the set of FFNs with arbitrary hidden layer size. Then we have:

1. For any finite measure μ on $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ and $p \in [1, \infty)$, the set $\mathcal{NN}_\infty$ is dense in $L^p(\mathbb{R}^d, \mu)$.
2. If further σ is continuous then $\mathcal{NN}_\infty$ is dense in $C(\mathbb{R}^d)$ in the sense that for any $f \in C(\mathbb{R}^d)$ there exists a sequence of neural networks $\varphi_n \in \mathcal{NN}_\infty$ that converges uniformly on any compact set $K \subset \mathbb{R}^d$ to f.

The Universal Approximation Theorem goes back to [Hornik et al., 1989].

Universal Approximation Theorem

Let σ be non-constant and bounded, $d \in \mathbb{N}$ and $\mathcal{NN}_\infty := \{\mathrm{FFN}_\theta^{d,L,1} \mid L \in \mathbb{N}\}$ be the set of FFNs with arbitrary hidden layer size. Then we have:

1. For any finite measure μ on $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ and $p \in [1, \infty)$, the set $\mathcal{NN}_\infty$ is dense in $L^p(\mathbb{R}^d, \mu)$.
2. If further σ is continuous then $\mathcal{NN}_\infty$ is dense in $C(\mathbb{R}^d)$ in the sense that for any $f \in C(\mathbb{R}^d)$ there exists a sequence of neural networks $\varphi_n \in \mathcal{NN}_\infty$ that converges uniformly on any compact set $K \subset \mathbb{R}^d$ to f.

Therefore, we may approximate any $\delta \in \mathcal{H}$ arbitrarily well using a FFN.

Heston Model

Definition (Heston Model)

In a Heston model, the dynamics of the underlying risky asset S are given by

$$dS_t = \mu S_t dt + \sqrt{V_t} S_t dB_t$$
$$dV_t = \alpha(b - V_t)dt + \sigma\sqrt{V_t} dW_t$$

where B, W are two correlated Brownian motions with correlation $\rho \in [-1, 1]$.

Heston Model

Definition (Heston Model)

$$dS_t = \mu S_t dt + \sqrt{V_t} S_t dB_t$$
$$dV_t = \alpha(b - V_t)dt + \sigma\sqrt{V_t} dW_t$$

In our case, we are modeling a 'typical equity market' using

$$S_0 = 100, \alpha = 1, b = V_0 = 0.04, \rho = -0.7, \mu = 0 \text{ and } \sigma = 2.$$

GBM vs Heston

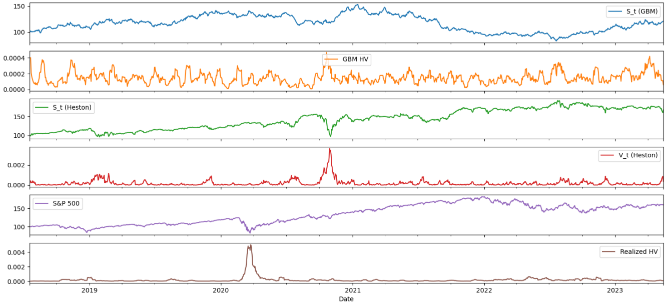

Figure 3: Simulated stock prices by GBM and Heston as well as the real S&P 500; Below each price we can see the corresponding (14 days rolling) volatility.

Idealized Variance Swap

In the paper, the hedging agent is allowed to trade in

- $S_t^{(1)}$, the underlying risky asset and
- an idealized variance swap which is defined as

$$
\begin{aligned}
S_t^{(2)} &:= \mathbb{E}_{\mathbb{Q}}\Big[\int_0^T V_s \, ds \,\Big|\, \mathcal{F}_t \Big] \\
&= \int_0^t V_s \, ds + \underbrace{\frac{V_t - b}{\alpha}(1 - e^{-\alpha(T-t)}) + b(T-t)}_{:=L(t,V_t)} .
\end{aligned}
$$

How this looks like

Figure 4: Histograms from 10^7 simulations of $S_T^{(1)}$ and $S_T^{(2)}$ in the Heston model for $T = 30$ and parameters as in the paper

How this looks like

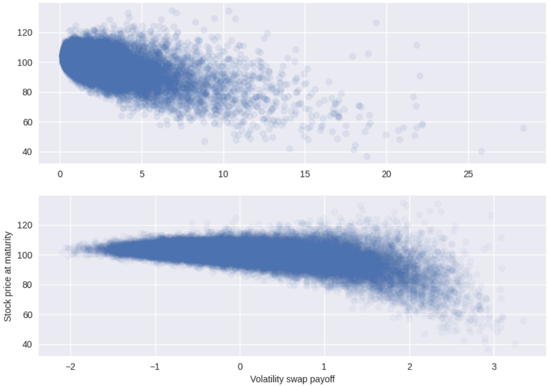

Figure 5: Scatter plot showing how $S_T^{(1)}$ and volatility swap payoff $S_T^{(2)}$ (or log thereof in the bottom plot) interact

How this looks like

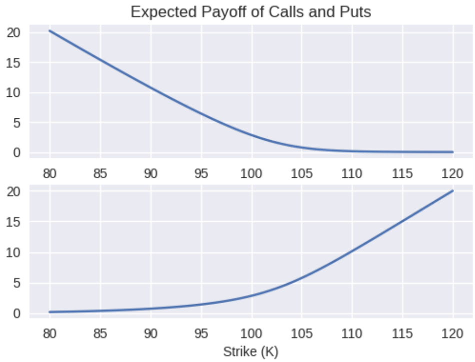

Figure 6: Expected Payoffs from Calls and Puts at $T = 30$ with strikes on the x-Axis; Assuming $r = 0$, these are the 'fair' option prices

How this looks like

Figure 7: Price of a variance swap throughout a hedging period

How this looks like

Figure 8: Price of a variance swap throughout a hedging period

Learning Hedging Strategies

Definition (Information Set)

Let I_t be the $\mathbb{F}$-adapted information set with all the information we want to provide our hedging agent with.

In our case, let

$$I_t := (S_t^{(1)}, S_t^{(2)})$$

since our market model is *markovian*.

Learning Hedging Strategies

Definition (Information Set)

Let I_t be the $\mathbb{F}$-adapted information set with all the information we want to provide our hedging agent with.

In our case, let

$$I_t := (S_t^{(1)}, S_t^{(2)})$$

since our market model is *markovian*.

We may also allow additional information to be included in I_{t-1}, like upcoming earnings or market sentiment if such information can be built into our model somehow.

Learning Hedging Strategies

Definition (Information Set)

Let I_t be the $\mathbb{F}$-adapted information set with all the information we want to provide our hedging agent with.

In our case, let

$$I_t := (S_t^{(1)}, S_t^{(2)})$$

since our market model is *markovian*.

When hedging in a *non-markovian* market, we might want to define

$$\tilde{I}_t := (S_t^{(1)}, S_t^{(2)}, S_{t-1}^{(1)}, S_{t-1}^{(2)}, \ldots, S_0^{(1)}, S_0^{(2)}).$$

Learning Hedging Strategies

Definition (Deep Hedging Strategy)

For $L \in \mathbb{N}, d \in \mathbb{N}$, $\sigma : \mathbb{R} \to \mathbb{R}$ an activation function, I_t our information set at time t and $\mathrm{FFN}_\theta^{d+k,L,d} : \mathbb{R}^{d+k} \to \mathbb{R}^d$, a feed forward neural network, we define our **deep hedging strategy** as

$$\begin{aligned} \delta_t^\theta :\ & \mathbb{R}^{d+k} \to \mathbb{R}^d \\ I_{t-1} \mapsto \delta_t^\theta(I_{t-1}) &:= \mathrm{FFN}_\theta^{d+k,L,d}(I_{t-1}) \\ &= W_2 \cdot \sigma(W_1 \cdot I_{t-1} + b_1) + b_2. \end{aligned}$$

Learning Hedging Strategies

Observe how our previous optimization problem

$$\pi(-Z) := \inf_{\delta \in \mathcal{H}} \rho(-Z + (\delta \cdot S)_T - C_T(\delta))$$

becomes, for $\delta^\theta \in \mathcal{H}$,

$$\pi(-Z) := \inf_{\theta \in \Theta} \rho(-Z + (\delta^\theta \cdot S)_T - C_T(\delta^\theta))$$

just the problem of finding the best parameters θ for our neural network.

Learning Hedging Strategies

$$\pi(-Z) := \inf_{\theta \in \Theta} \rho(-Z + (\delta^\theta \cdot S)_T - C_T(\delta^\theta))$$

Further, if ρ and c_t are differentiable, we can compute

$$\nabla_\theta \rho(-Z + (\delta^\theta \cdot S)_T - C_T(\delta^\theta))$$

in order to update our parameters θ.

The case for OCE Risk Measures

If $\rho = \rho_l$ is an OCE risk measure, then the problem becomes:

$$\begin{aligned}
\pi(-Z) &= \inf_{\theta\in\Theta} \inf_{w\in\mathbb{R}} w + \mathbb{E}[l(Z - (\delta^\theta \cdot S)_T + C_T(\delta^\theta) - w)] && (1)\\
&= \inf_{\tilde{\theta}\in\tilde{\Theta}} w + \mathbb{E}[l(Z - (\delta^\theta \cdot S)_T + C_T(\delta^\theta) - w)] && (2)\\
&= \inf_{\tilde{\theta}\in\tilde{\Theta}} J(\tilde{\theta}) && (3)
\end{aligned}$$

For $\tilde{\Theta} := \Theta \times \mathbb{R}$ and of course $\tilde{\theta} := (\theta, w)$.

The case for OCE Risk Measures

If $\rho = \rho_l$ is an OCE risk measure, then the problem becomes:

$$\pi(-Z) = \inf_{\theta \in \Theta} \inf_{w \in \mathbb{R}} w + \mathbb{E}[l(Z - (\delta^\theta \cdot S)_T + C_T(\delta^\theta) - w)] \tag{1}$$

$$= \inf_{\tilde{\theta} \in \tilde{\Theta}} w + \mathbb{E}[l(Z - (\delta^\theta \cdot S)_T + C_T(\delta^\theta) - w)] \tag{2}$$

$$= \inf_{\tilde{\theta} \in \tilde{\Theta}} J(\tilde{\theta}) \tag{3}$$

We can now compute the gradient of J with respect to $\tilde{\theta}$ and use stochastic gradient descent to find the optimal parameters $\hat{\theta}$.

The case for OCE Risk Measures

If $\rho = \rho_l$ is an OCE risk measure, then the problem becomes:

$$\pi(-Z) = \inf_{\theta \in \Theta} \inf_{w \in \mathbb{R}} w + \mathbb{E}[l(Z - (\delta^\theta \cdot S)_T + C_T(\delta^\theta) - w)] \tag{1}$$

$$= \inf_{\tilde{\theta} \in \tilde{\Theta}} w + \mathbb{E}[l(Z - (\delta^\theta \cdot S)_T + C_T(\delta^\theta) - w)] \tag{2}$$

$$= \inf_{\tilde{\theta} \in \tilde{\Theta}} J(\tilde{\theta}) \tag{3}$$

Specifically, we sample $\omega_1, \ldots, \omega_{N_{\text{batch}}} \in \Omega$, estimate

$$\tilde{J}(\tilde{\theta}) = w + \sum_{m}^{N_{\text{batch}}} l(Z(\omega_m) - (\delta^\theta \cdot S)_T(\omega_m) + C_T(\delta^\theta)(\omega_m) - w) \frac{N \cdot \mathbb{P}(\{\omega_m\})}{N_{\text{batch}}}$$

from which we immediately get $\nabla_{\tilde{\theta}} \tilde{J}(\tilde{\theta})$ using autodiff.

Neural Network Architecture in the Paper

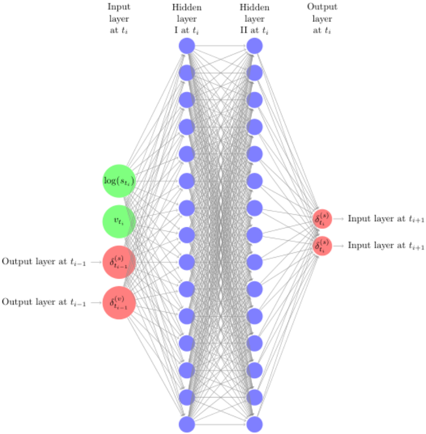

Figure 9: (Recurrent) Neural Network Architecture used in the paper

Learned and Model Delta

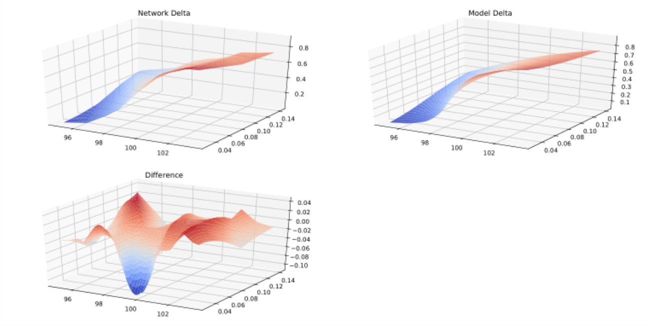

Figure 10: Learned $\delta^{\theta,1}$ (the position in the underlying) and Heston model delta δ as a function of $(S_t^{(1)}, V_t)$ for $t = 15$ days.

Key Takeaways

- Deep hedging is a **model-independent** approach to hedging which can be applied to any market model.
- It is **computationally efficient** and can be used to hedge in **high-dimensional** (i.e. multi-asset) markets.
- It is **flexible** in the sense that we can choose
  - the risk measure we want to optimize for,
  - transaction cost and trading restrictions,
  - the information set we provide to the agent and
  - the hedging instruments we allow trading in.
- The approach is **easy to adapt** to more sophisticated derivatives or to the hedging of multiple liabilities simultaneously.

Criticism of the Paper

- Optimization via $\nabla_{\tilde{\theta}}(\tilde{J})(\tilde{\theta})$ directly is actually really hard and rather primitive compared to more modern reinforcement learning approaches.

Thank you for your attention!

References

Bühler, H., Gonon, L., Teichmann, J., and Wood, B. (2018).
Deep hedging.

Hornik, K., Stinchcombe, M., and White, H. (1989).
Multilayer feedforward networks are universal approximators.
Neural networks, 2(5):359–366.