

# Better Hedging with Reinforcement Learning

전통적 Black-Scholes 모델의 한계를 극복하고 AI를 활용한 혁신적인 헤징 전략을 탐색합니다



## 목차

1. 전통적 헤징의 도전과제
2. 변동성 스마일과 시장 현실
3. AI와 강화학습의 등장
4. Deep Hedging의 개념
5. RL 헤징 아키텍처- AAPL 예시
6. 성과 비교 - 전통적 vs AI 헤징

# 1. 전통적 헤징의 도전과제 - Black-Scholes 모델의 한계



## 변동성 일정성 가정

BSM은 자산 가격의 변동성이 일정하다고 가정하지만, 실제 시장에서는 변동성이 시간과 자산 가격에 따라 크게 변동합니다. 이것이 변동성 스마일/스커 현상을 설명하지 못하는 근본적 원인입니다.

실제 시장에서 관찰되는 변동성 스마일 현상 - BSM 모델로 설명 불가



## 거래비용 및 유동성 무시

BSM은 거래비용이 없다고 가정하여 완벽한 헤징이 가능하다고 전제합니다. 그러나 현실에서는 거래수수료, 스프레드, 시장충격 비용이 발생하여 이론적인 헤징 전략의 성과를 크게 훼손합니다.



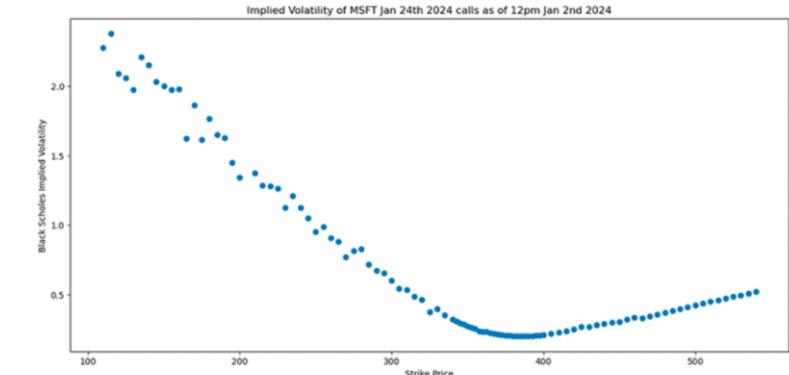
## 유럽형 옵션 제한

기본 BSM 모델은 유럽형 옵션(만기일에만 행사 가능)에만 적용 가능합니다. 미국형 옵션이나 이색옵션(Exotic Options)에는 복잡한 수정이 필요합니다.



## 배당 및 금리 가정

배당금이 없고 금리가 일정하다는 가정은 현실과 괴리가 있으며, 특히 장기 옵션 평가시 더 큰 문제가 됩니다.



**Figure 7.1** The volatility “smile” and “skew”: implied volatility of MSFT call options expiring on January 24, 2024, as of 12:00 pm on January 2, 2024. MSFT stock price was about \$375/share. Implied volatility calculated using BSM from the mid-prices of call options.

Black-Scholes 모델은 50년 가까이 사용되었지만, 현대 시장의 복잡성을 완전히 반영하지 못합니다.

## 2. 변동성 스마일과 현실적 문제들 - 시장 현실



### 변동성 스마일 현상

외가격(OTM)과 내가격(ITM) 옵션의 내재 변동성이 등가격(ATM) 옵션보다 높게 나타나는 현상으로, 스트라이크 가격에 따라 U자형 곡선을 형성합니다.



### Black-Scholes 모델의 한계

BSM 모델은 일정한 변동성을 가정하므로 변동성 스마일을 설명하지 못합니다. 이 모델에 따르면 변동성 곡선은 평평해야 하지만, 실제 시장은 그렇지 않습니다.



### 옵션 가격 왜곡

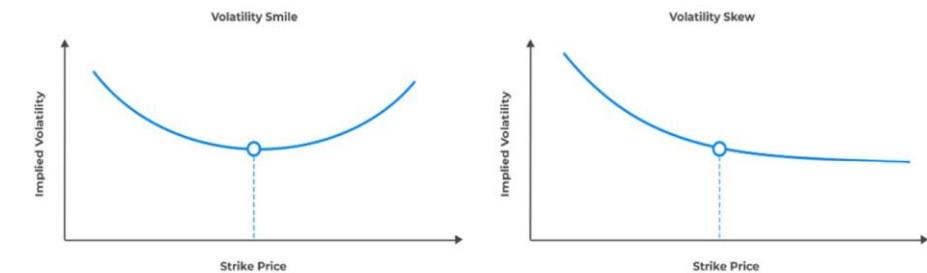
시장 스트레스 상황에서 OTM 풋옵션의 내재 변동성이 급등하는 '변동성 스커' 현상이 발생합니다. 이는 시장의 꼬리 위험(tail risk)에 대한 투자자들의 우려를 반영합니다.



변동성 스마일은 1987년 블랙 먼데이 이후 더욱 두드러지게 관찰되기 시작했으며, 이는 시장 붕괴에 대한 위험 프리미엄을 반영합니다.



Volatility Skew vs Volatility Smile



현실 세계에서 옵션 가격은 수요와 공급 등 다양한 요인의 영향을 받을 수 있습니다. BSM 모델이 옵션 가격의 변동성을 완전히 반영하지 못함을 나타냅니다.

### 3. AI의 등장 - 왜 강화학습인가?



#### 머신러닝/강화학습(ML/RL) 발전 동향

최근 딥러닝과 강화학습 기술의 획기적인 발전으로 금융 분야에서 복잡한 의사결정 문제를 해결할 수 있게 되었습니다. 특히 2020년 이후 대형 투자은행들의 AI 기술 도입이 급속히 증가하고 있습니다.



#### 데이터 기반의 동적 적응력

강화학습은 실시간 시장 데이터를 학습하고 적응하는 능력이 탁월합니다. 시장 상황과 변동성이 변화해도 지속적으로 학습하고 전략을 최적화하여 전통적인 모델의 정적 접근 방식 한계를 극복합니다.



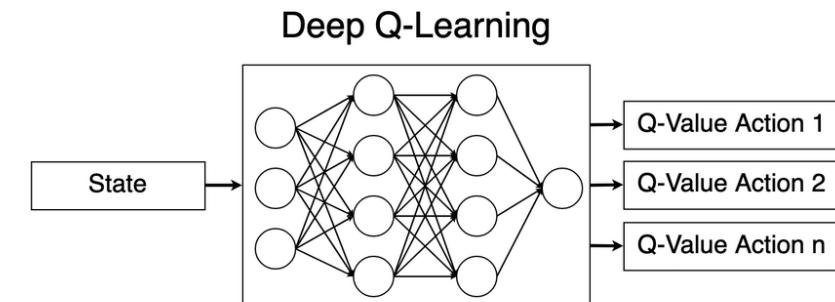
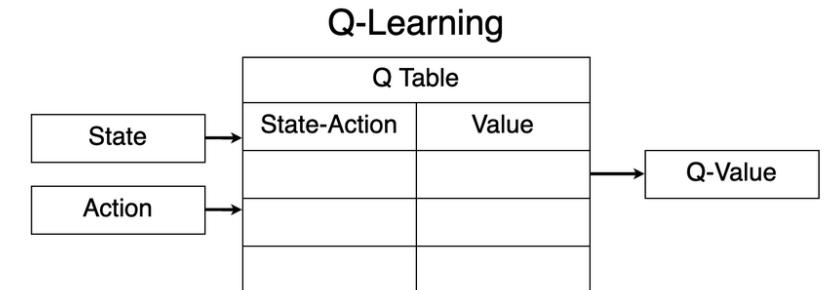
#### 시장 변화 및 거래비용 반영

거래비용, 유동성 제약, 시장 충격 등 현실적인 시장 조건을 모델에 통합할 수 있어 실전 환경에서의 성능이 뛰어납니다. 전통적인 Black-Scholes 접근법과 달리 현실 세계의 마찰 요인을 고려합니다.



#### 비선형·복잡 보상구조 처리

강화학습은 비선형적이고 복잡한 보상 구조를 자연스럽게 처리할 수 있어, 다양한 리스크 선호도와 수익 최적화 목표를 모델링하는데 탁월합니다. 이는 특히 복잡한 파생상품 헤징에 유리합니다.



강화학습의 환경-에이전트-보상 프레임워크

## 4. Deep Hedging 개념과 원리 - AI 기반 헤징의 혁신



### 신경망과 강화학습 기반 전략

Deep Hedging은 복잡한 신경망과 강화학습 알고리즘을 활용해 최적의 헤징 전략을 학습합니다. 옵션의 정당한 가격이나 근사적인 복제전략의 가정 없이 직접적으로 헤징 결정을 최적화합니다.



### 빅데이터와 실시간 정보 활용

방대한 시장 데이터를 학습하여 기존 모델이 설명할 수 없는 시장 패턴을 포착합니다. 과거 데이터뿐 아니라 실시간 정보를 활용해 시장 변화에 즉각적으로 대응할 수 있습니다.



### 리스크 측정법 최적화

전통적인 델타-중립 헤징과 달리, 다양한 리스크 측정 지표(CVaR, 기대효용 등)를 직접 최적화할 수 있습니다. 이를 통해 투자자의 리스크 성향에 맞춘 맞춤형 헤징이 가능합니다.



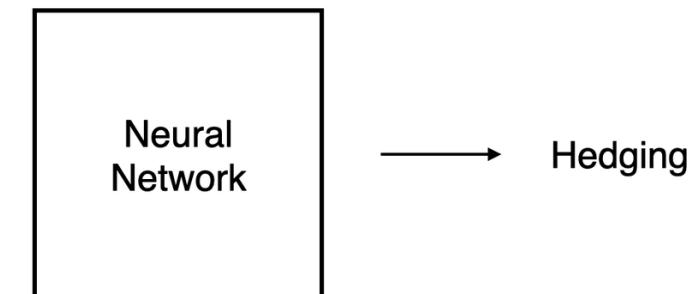
### 거래비용 및 시장 마찰 고려

거래비용, 유동성 제약, 시장충격 등 현실적 제약조건을 직접 모델에 반영하여 이론적으로 완벽하지만 실현 불가능한 전략이 아닌, 실제 실행 가능한 최적 전략을 도출합니다.

"Deep Hedging은 Black-Scholes 가정에 의존하지 않고, 시장 데이터에서 직접 학습하여 최적의 헤징 전략을 발견합니다."

### Deep Hedging 신경망 아키텍처

Market Information → Neural Network



입력 레이어  
시장 데이터, 옵션  
특성, 시간 정보, 헤지  
포지션 상태 등 관련  
금융 데이터

은닉 레이어  
복잡한 비선형 패턴을  
포착하는 다중 레이어  
네트워크. LSTM이나  
GRU 구조로 시계열  
의존성 포착

출력 레이어  
최적의 헤징 행동(델타)  
결정. 시장 변화에 맞춰  
동적으로 포지션  
조정량 산출

## 5. RL 헤징 아키텍처 - AAPL 예시



### 강화학습 기본 메커니즘

에이전트(헤징 전략)가 환경(금융시장)과 상호작용하며 보상(리스크 감소)을 최대화하는 과정에서 최적의 헤징 전략을 학습합니다. 시장 변화에 동적으로 적응할 수 있는 딥러닝 기반의 접근법입니다.



### RL 헤징 핵심 구성요소

- 상태 (State): 시장 가격, 변동성, 포지션, 시간 등
- 행동 (Action): 헤지 포지션 조정량
- 보상 (Reward): 리스크 감소 - 거래비용



### 파이프라인(4단계)

1. 식별: 기초자산의 가격과정을 정해서 모수를 잡음
2. 시뮬레이션 훈련: 이론 무차의 관계를 따르는 모의데이터로 기본기를 학습
3. 실데이터 개선: 최근 시장 데이터로 정책을 미세조정
4. 테스트/구현: 표본외 성능 감시하며 운영·개선

### 예시 프로젝트 목적

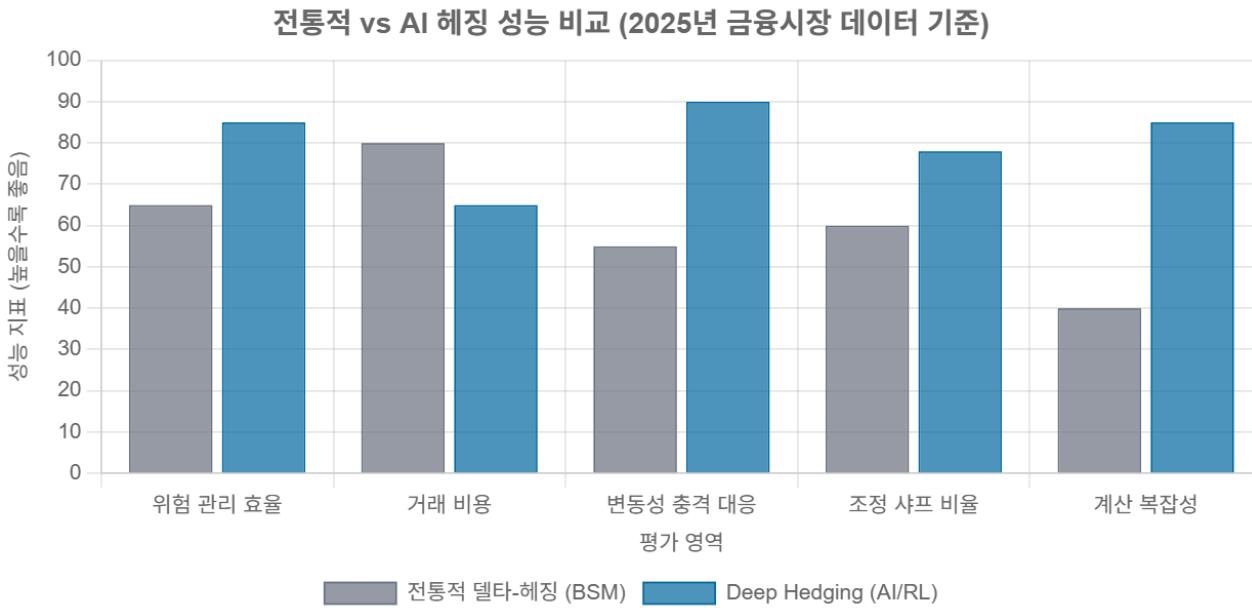
해지 대상 자산(AAPL 콜 옵션)과 헤징 수단(이 경우 기초 주식)을 포함하는 해지된 포트폴리오의 손익(PnL) 분산을 최소화하는 에이전트 또는 최적의 헤징 정책을 얻는 것이다.

### 확률적 정책(가우시안) 기반의 정책경사(policy-gradient) + 모방학습

- 정책: 네트워크가  $(\mu, \sigma)$ 를 출력  $\rightarrow$  다음 헤지 포지션을 확률적으로 정함.
- 보상/손실: 매 스텝 PnL 변화에서 손실만 페널티, 거래비용 포함.
- 학습 흐름
  1. 모방학습으로 BS 델타에 가깝게 시작(초기 정책 안정화)
  2. 정책경사로 실제 PnL 페널티 최소화 방향으로 미세조정  
 $\rightarrow$  결국 “돈을 덜 잃고, 쓸데없는 리밸런싱을 줄이는” 방향으로 확률을 재배치

\*언제 정책경사를 쓰면 좋은가?  
 $\rightarrow$  행동이 연속형일 때(헤지 수량, 가격, 속도 등)
- 정책경사 = “좋았던 행동의 확률은 키우고, 나빴던 건 줄이는” 확률 조정법.
- 가우시안 정책을 쓰면 연속형 헤지 같은 문제에 깔끔하게 들어맞는다.

## 6. 성과 비교 - 전통적 vs AI 헤징



리스크 감소율 +32%

32% 향상

변동성 환경에서 최대 손실 감소

실행 시간 -68%

5배 빠름

복잡한 상품 헤징 시 계산 속도

시스템 적응력 +85%

시장 충격 감내

극단적 시장 상황 대응 능력

### 헤징 오차 감소

Deep Hedging은 거래 비용과 시장 마찰을 고려해 평균 26% 더 낮은 헤징 오차를 보이며, 특히 변동성이 급변하는 환경에서 우수한 성능을 입증했습니다.

### 거래 비용 최적화

AI 헤징은 거래 비용을 인식하고 최적의 거래 타이밍과 수량을 결정하여 전통적 델타 헤징 대비 총 비용을 18% 절감합니다.

### 실시간 시장 적응

강화학습 모델은 변화하는 시장 조건에 동적으로 적응하여 스트레스 상황이나 변동성 급증 시에도 일관된 헤징 성능을 유지합니다.

### 실증 연구 결과

2025년 최신 연구에 따르면, Deep Hedging 전략은 모든 시장 사이클에서 샤프 비율이 0.32 향상되었으며, 최대 손실폭이 28% 감소했습니다.