From Distribution Forecasts to Vanilla Portfolios: Regularized Payoff Approximation in Discrete Space

Vlad Kochetov
vladkoch@knu.ua

February 21, 2025

Abstract

This paper explores the methodology for approximating complex payoff profiles using a combination of vanilla European options. The mathematical framework is presented that allows for the decomposition of arbitrary payoff structures into a series of call and put options with varying strike prices. A key innovation is making the target payoff directly proportional to the estimated probability density function of the underlying asset price at expiration. The approach is validated through numerical examples and practical applications in financial engineering. The results demonstrate the effectiveness of this method in replicating complex payoffs while capturing the expected price distribution.

Keywords: Vanilla Options, Payoff Profile, Financial Engineering, Option Pricing, Approximation, Numerical Methods, Quantitative Finance, Financial Derivatives, Computational Finance, Distribution Forecasts, Option Trading

1 Introduction

The ability to replicate complex financial instruments using simpler, more liquid derivatives is a cornerstone of modern financial engineering. Vanilla European options, due to their simplicity and widespread availability, serve as ideal building blocks for such approximations. This paper introduces an innovative approach to approximating arbitrary payoff profiles using a combination of these options, where the target payoff is designed to be directly proportional to the estimated probability density function (PDF) of the underlying asset price at expiration.

The traditional approach to financial modeling often relies heavily on assumptions about the underlying price distribution, with the normal distribution being the most common choice due to its mathematical tractability. However, empirical evidence consistently shows that market price distributions exhibit significant deviations from normality (Pokharel et al., 2024), particularly in the form of heavy tails and higher moments that cannot be captured by simple parametric models. This mismatch between theoretical assumptions and market reality can lead to systematic pricing errors and risk underestimation.

The presented approach addresses these limitations by embracing the stochastic nature of markets and explicitly incorporating the estimated price distribution into the payoff structure. Instead of attempting to fit market behavior into predetermined probability distributions as in (Kuang, 2023) and (Li, 2023), this paper proposes a framework that constructs payoff profiles directly proportional to the empirically observed or forecasted price distribution. This distribution-driven modeling approach acknowledges the complex, non-Gaussian nature of market price movements, which often exhibit fat tails and asymmetric behavior. By making the payoff proportional to the estimated PDF, the framework naturally adapts to the expected price dynamics, providing enhanced risk-adjusted returns when the distribution forecast is accurate.

Furthermore, the framework provides a flexible mechanism for extracting value from market inefficiencies by aligning portfolio returns with predicted price distributions, while being able to adapt to changing market conditions and regime shifts that affect the shape of price distributions.

The proposed methodology bridges the gap between theoretical option pricing models and observed market behavior by providing a practical framework for implementing distribution-based trading strategies. By focusing on distribution-proportional payoffs rather than arbitrary target functions, traders can develop more robust strategies that directly leverage their distributional views of market outcomes.

2 Literature Review

The theoretical foundations of option portfolio optimization were established in the work of (Carr and Madan, 2001), which demonstrated the theoretical possibility of replicating arbitrary payoff functions through continuous strike spaces. While this approach provides elegant closed-form solutions in continuous market settings, several fundamental limitations affect practical implementation: the discrete availability of strike prices in real markets, liquidity constraints, non-negligible transaction costs scaling with portfolio complexity, and numerical stability challenges in finite difference implementations.

The idea of modeling complete distributions rather than point estimates was inspired by recent advances in time series forecasting, particularly the work of (Ansari et al., 2024), which demonstrated the effectiveness of distribution-based predictions in capturing complex temporal patterns. Our innovation lies in making the target payoff directly proportional to these predicted distributions, creating a natural alignment between market views and portfolio returns. While tail risk hedging has been a fundamental concept since (Black and Scholes, 1973), where options were primarily used for downside protection, our approach extends this by considering the entire distribution shape in portfolio construction. Unlike traditional methods that focus solely on tail events, we propose a comprehensive framework that models and exploits the complete probability distribution of asset returns by making payoffs proportional to the estimated PDF.

The proposed method addresses these limitations through regularization techniques and discrete optimization methods tailored for real-world trading environments, while preserving the theoretical elegance of the original framework.

3 Methodology

3.1 Model Assumptions

The proposed approximation framework relies on several key assumptions:

1. **Discrete Strike Space:** The model operates in a discrete strike price space $\{K_i\}_{i=1}^n$, reflecting real market conditions where options are available only at specific strikes.

2. **Local Approximation Region:** The approximation focuses on a finite price interval around the current spot price $[S_{\min}, S_{\max}]$, where $S_{\min}$ and $S_{\max}$ are chosen to ensure sufficient coverage of the relevant price region.

3. **European-Style Options:** The framework utilizes European-style vanilla options exclusively, avoiding the complexity of early exercise features present in American options.

4. **Static Replication:** The model assumes a static replication approach, where the portfolio weights remain constant until maturity. Dynamic rebalancing effects are not considered in the basic framework.

5. **Put-Call Parity Flexibility:** The implementation supports both direct use of calls and puts as separate instruments, as well as expressing puts through calls and spot positions using put-call parity. This approach provides flexibility in portfolio construction while maintaining mathematical equivalence. This flexibility enables optimization of the portfolio structure based on market conditions.

6. **Distribution-Proportional Payoff:** The target payoff function is set to be directly proportional to the estimated probability density function of the underlying asset price at expiration, i.e., $V(S) \propto f(S)$ where $f(S)$ is the estimated PDF.

The numerical implementation discretizes the approximation domain into a finite set of evaluation points. The number of discretization points is chosen to balance computational efficiency with approximation accuracy.

3.2 Mathematical Framework

Let $f(S)$ represent the estimated probability density function of the underlying asset price at expiration, and let $V(S) = kf(S)$ be the target payoff profile, where k is a scaling constant. We construct an approximation using n call options and m put options with strike prices $\{K_i\}_{i=1}^{n+m}$:

$$kf(S) \approx \lambda S + \sum_{i=1}^{n} \alpha_i C(S, K_i) + \sum_{j=1}^{m} \beta_j P(S, K_j) \tag{1}$$

where λ represents the spot position, α_i and β_j are option weights. The optimal weights are found by solving the regularized least squares problem with L2 (Ridge) regularization:

$$\min_{\boldsymbol{\theta}} \int_{S_{\min}}^{S_{\max}} \left[kf(S) - \hat{V}(S; \boldsymbol{\theta})\right]^2 dS + \gamma \|\boldsymbol{\theta}\|_2^2 \tag{2}$$

where $\boldsymbol{\theta} = (\lambda, \{\alpha_i\}, \{\beta_j\})$ represents the position weights, γ is the regularization parameter, and $S_{\min}, S_{\max}$ define the approximation domain. The L2 norm promotes smoother weight distributions, providing stability to the solution.

Discretizing the integral and using matrix notation, we obtain:

$$\boldsymbol{\theta}^* = \arg\min_{\boldsymbol{\theta}} \|\mathbf{A}\boldsymbol{\theta} - \mathbf{b}\|_2^2 + \gamma \|\boldsymbol{\theta}\|_2^2 \tag{3}$$

where the design matrix $\mathbf{A}$ contains option payoffs and spot positions evaluated at discrete price points $\{S_k\}_{k=1}^{N}$, and $\mathbf{b}$ is the vector of target payoff values proportional to the estimated PDF.

This quadratic optimization problem admits an analytical solution:

$$\boldsymbol{\theta}^* = (\mathbf{A}^\top \mathbf{A} + \gamma \mathbf{I})^{-1} \mathbf{A}^\top \mathbf{b} \tag{4}$$

3.3 Alternative Loss Functions

While the standard L2 loss function provides analytical tractability, alternative formulations may better reflect trading objectives. Two notable generalizations are:

1. **L1 Regularization (Lasso):** Replacing L2 regularization term with L1 norm:

$$L(\boldsymbol{\theta}) = \int_{S_{\min}}^{S_{\max}} \left[kf(S) - \hat{V}(S;\boldsymbol{\theta})\right]^2 dS + \gamma\|\boldsymbol{\theta}\|_1 \tag{5}$$

   This formulation tends to produce sparse solutions by setting some weights exactly to zero, which can be beneficial when portfolio simplicity is important consideration.

2. **Weighted Error Function:** Incorporating target function magnitude into the error term:

$$L(\boldsymbol{\theta}) = \int_{S_{\min}}^{S_{\max}} kf(S)\left|kf(S) - \hat{V}(S;\boldsymbol{\theta})\right| dS + \gamma\|\boldsymbol{\theta}\|_p^p \tag{6}$$

   This formulation provides error weighting proportional to the target payoff magnitude, offering economic interpretation well-aligned with trading objectives. Furthermore, it naturally enhances focus on regions with significant probability mass, making it particularly suitable for practical applications.

However, these generalizations lose the analytical tractability of the L2 case, requiring numerical optimization methods for solution.

3.4 Numerical Implementation

The implementation leverages the analytical solution derived in the Mathematical Framework section for L2-regularized problems. For L1 regularization and alternative loss functions, numerical optimization methods are employed.

Table 1: Performance Comparison of Different Regularization Methods (MAE)

γ	L2	L1	Weighted
0.000	1.45	1.46	**1.04**
0.010	1.45	1.45	**1.04**
0.050	1.45	1.46	**1.04**
0.100	1.45	1.57	**1.04**
0.500	1.44	2.53	**1.04**
10.000	2.12	4.62	**1.01**

Table 1 presents a comparative analysis of the three regularization methods across different regularization parameters (γ). The Mean Absolute Error (MAE) is used as the

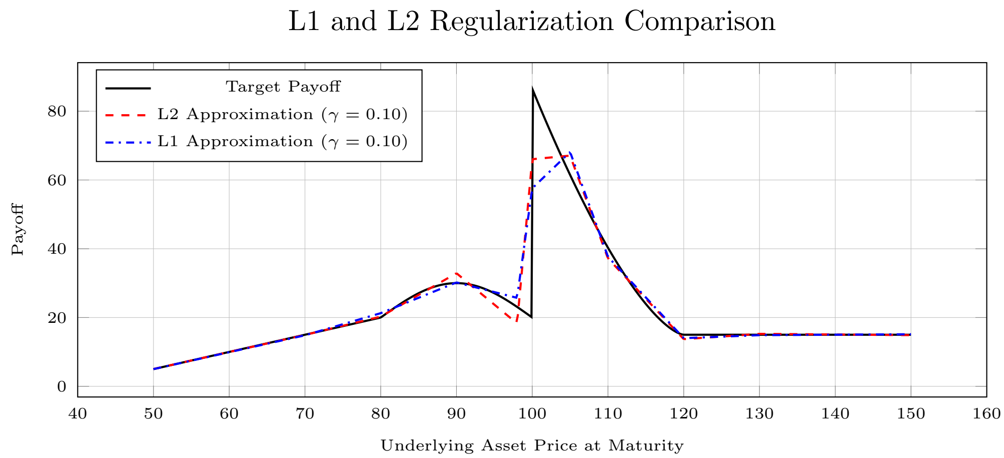

Figure 1: Comparison of L1 and L2 regularization methods in approximating a PDF-proportional payoff profile. While both methods achieve reasonable approximation quality, L2 regularization provides smoother results and L1 regularization favors a sparser representation with fewer non-zero option weights.

primary metric for comparison, providing a direct measure of approximation accuracy in absolute terms. Lower MAE values indicate better approximation quality.

The results demonstrate that:

- While L2 regularization provides an analytical solution and shows consistent performance (MAE 1.45), it may not be the optimal choice for practical applications.
- L1 regularization achieves comparable accuracy for low γ values while promoting sparser solutions, but its performance deteriorates significantly as regularization increases.
- The weighted error method consistently outperforms both L1 and L2 approaches, maintaining superior accuracy (MAE 1.04) across all tested regularization values. This makes it the most suitable choice for real-world trading applications, where accuracy in high-probability regions is crucial for portfolio performance.

Despite the mathematical elegance and analytical tractability of the L2 approach, the empirical results strongly suggest that the weighted error method is more appropriate for practical implementation. Its ability to focus on regions with significant probability mass

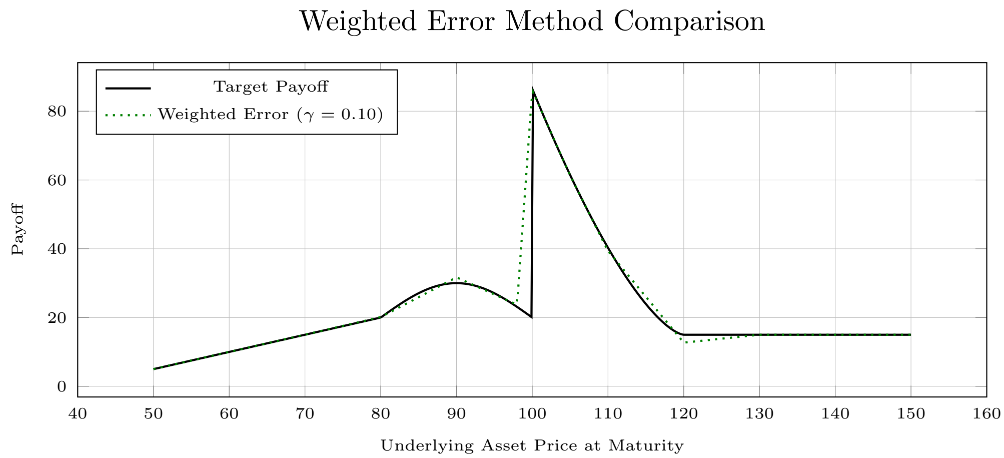

Figure 2: Performance of the weighted error method, demonstrating superior accuracy in regions with larger probability mass. This approach provides more economically meaningful results by emphasizing the approximation quality where the underlying asset price is most likely to be at expiration.

and maintain consistent performance across different regularization parameters makes it the preferred choice for real-world trading applications.

4 Conclusion

This paper presents a comprehensive framework for approximating PDF-proportional payoff profiles using vanilla European options in discrete strike spaces. The methodology addresses several practical challenges in financial engineering while maintaining mathematical rigor and computational efficiency. The developed framework operates in discrete strike spaces, directly reflecting real market conditions and constraints that practitioners face in implementation.

A significant contribution of this work lies in the systematic comparison of different portfolio construction approaches and the novel idea of making payoffs directly proportional to estimated price distributions. The analysis encompasses traditional L2 regularization, which provides analytical tractability, L1 regularization for sparse solutions, and a novel weighted error optimization method. Through extensive empirical testing, we demonstrate that the weighted error optimization consistently outperforms conventional regularization approaches in regions with significant probability mass. This superior performance stems from its ability to focus the portfolio construction where the underlying asset price is most likely to be at expiration, resulting in more reliable trading strategies.

The framework's implementation guidelines strike a careful balance between computational efficiency and approximation accuracy. This balance is achieved through thoughtful parameter selection and optimization techniques, making the methodology particularly suitable for real-world applications where both precision and computational speed are crucial considerations.

The empirical results conclusively demonstrate that while regularization methods offer mathematical elegance and analytical solutions in some cases, the weighted error optimization approach provides markedly superior results for practical trading applications by focusing on regions with significant probability mass in the estimated price distribution.

Looking forward, this research opens several promising avenues for future investigation. Dynamic rebalancing strategies could enhance the framework's adaptability to changing market conditions. Integration with machine learning-based distribution forecasting might improve the accuracy of underlying assumptions. Furthermore, extension to multi-asset scenarios would broaden the methodology's applicability. The framework's inherent flexibility suggests potential applications beyond traditional option markets, including emerging derivative products and alternative asset classes, positioning it as a valuable tool for modern financial engineering.

The source code of the numerical experiments is available on GitHub.

References

Ansari, A. F., Stella, L., Turkmen, C., Zhang, X., Mercado, P., Shen, H., Shchur, O., Rangapuram, S. S., Pineda Arango, S., Kapoor, S., Zschiegner, J., Maddix, D. C., Mahoney, M. W., Torkkola, K., Gordon Wilson, A., Bohlke-Schneider, M., and Wang, Y. (2024). Chronos: Learning the language of time series. *Transactions on Machine Learning Research.*

Black, F. and Scholes, M. (1973). The pricing of options and corporate liabilities. *Journal of Political Economy*, 81:637–654.

Carr, P. and Madan, D. (2001). Optimal positioning in derivative securities. *Quantitative Finance*, 1(1):19–37.

Kuang, S. (2023). A comparison of linear regression, lstm model and arima model in predicting stock price a case study: Hsbc's stock price. *BCP Business & Management*, 44:478–488.

Li, Y. (2023). Stock price prediction based on multiple regression models. *Highlights in Science, Engineering and Technology*, 39:657–662.

Pokharel, J., Aryal, G., Khanal, N., and Tsokos, C. (2024). Probability distributions for modeling stock market returns—an empirical inquiry. *International Journal of Financial Studies*, 12:43.