Price and Hedge Discovery of Bermudan Swaptions under Hull-White Dynamics using Q-Learning

Samarth Kadaba

SMBC Capital Markets, New York, NY

August 24, 2023

About Me

I am Electrical Engineering student at Stanford University broadly interested in optimization and control theory.

In my free time, I do research in robotics and dance on my schools Bhangra team!

Figure: Enjoying New York City with friends.

Goals and Achievements

Goals

1. Develop understanding of...
   - Classical option pricing theory
   - Linear interest-rate derivatives (IRDs)
   - Term structure modeling
2. Understand DP-based and data-driven Q-Learning
3. Connect (1) and (2) for **model-free** IRD pricing

Achievements

1. Optimal hedges/prices discovered with Q-learning
2. Demonstrated algorithm tractability and model customization

Problem

Objective: Determine optimal price of Bermudan swaption using underlying swap trajectories (S_t) and replicating portfolio Π_t with terminal payoffs $\Pi_T = H_T(S_T)$ [1]

1. Preliminaries on Q-Learning in a Black-Scholes-Merton word (QLBS)
2. Considerations for interest-rate derivative products
3. Methods for pricing/hedging European and Bermudan swaptions
4. Selected results

[1] I. Halperin (2019). *QLBS: Q-Learner in the Black-Scholes(-Merton) Worlds*. arXiv: 1712.04609 [q-fin.CP]

Preliminaries on Q-Learning

Benefits of Q-Learning

- Provides optimal price **and** hedge actions
- Can be relaxed: 1) non-constant volatility, 2) incorporate dividends, etc. via reward structure
- Converges to BSM hedges and price as $\Delta t \to 0$ (derivation)

(Simplified) RL Approach via Maximization:

$$\begin{aligned} Q_t^* = \max_{a_t \in A} \quad & \mathbb{E}_t[R_t + \gamma \max_{a_{t+1}} Q_{t+1}^*] \\ \text{s.t.} \quad & \Pi_T = H_T(S_T) \end{aligned} \tag{1}$$

Condition on previous state, action value. R_t a function of -Var(Π_t) and current portfolio return.

DP Solution for QLBS (focus of methods herein)

QLBS aims to build a table of (state, action) tuples which learns optimal hedges for a given state (S_t), an **on-policy** learner.

Solve for Q values backwards in time with $-Q_0^*$ being the option price. Why is this the case?

- $-Q_0^*$, by construction, will be (expectation of) Π_0 plus some constant times $Var(\Pi_{t=0,\ldots,T})$.
- a_t^* found to minimize sum of negative returns and portfolio variance
- Fix a_t^*, find Q_t^* to minimize difference to instantaneous rewards plus discounted future rewards (2)

Solution via Fitted Q-Learning

In this setting, do not consider reward or state dynamics to be known, hence **model-free** and **off-policy**.

Consider Q value updates from single step observations $(X_t^n, a_t^n, R_t^n, X_{t+1}^n)$

From Bellman Optimality, solve **Least Squares** problem:

$$\min \quad \mathbb{L}_t(Q_t) = \sum_{k=1}^{N_{MC}} \left(R_t(a_t^*) + \gamma \max_{a_{t+1}} Q_{t+1}^*(a_{t+1}) - Q_t^*(a_t) \right)^2 \qquad (2)$$

Adaptation to Interest-Rate derivative Products

Added complexity from interest rate dynamics:

1. Construction of yield curve from market instruments
2. Evolution of term structure via short-rate model
3. Pricing for
   - Zero-Coupon Bonds
   - Swaps
   - European swaptions
   - Bermudan swaptions

Added complexity from early exercise:

1. Model as basket of European swaptions
2. Computation of continuation-values (cost of exercising)

Yield Curve Construction and Short-Rate Evolution

Term-structure fitted to current-market data, then evolved with HW.

1. Build **yield curve** from convexity-adjusted futures (forward-Libor and forward-swap rates).
2. HW (3) models the short-rate dynamics under the **risk-neutral** measure after calibration to market-quoted swaption volatilities.

$$dr(t) = [\theta(t) - \alpha(t)r(t)]dt + \sigma(t)dW(t) \tag{3}$$

We consider a simplified (1-factor) HW model with constant volatility, mean-reversion.

Pricing Swaps and European Swaptions

Swap valuation formula (4) gives present value of a fixed and floating-leg bonds according to specified payment (tenor) structure.

$$V_{swap}(t) = \sum_{n=0}^{N-1} \left(P(t, T_n) - P(t, T_{n+1}) - \tau_n k P(t, T_{n+1}) \right) \qquad (4)$$

- $P(t, T)$ denotes zero coupon bond price at time t with maturity T
- k denotes fixed rate (value which makes $V_{swap}(0) = 0$)
- $n = 0 ... N - 1$ denotes the N payment periods on the bond legs

Then, $V_{swaption}(t)$ with expiry T_0 is given by is simply the positive part of the swap value at time T_0 deflated by money-market account.

Pricing Bermudans in Longstaff Schwartz

The price of a bermudan at time T is the maximum of its intrinsic value and its computed continuation value at T (5) [2].

The continuation value $C(T)$ is given by regressing future Bermudan values on basis functions (typically of swap rates).

$$P(T) = max(P(T), C(T)) \tag{5}$$

Where $P(T)$ denotes the time T intrinsic value of the bermudan (i.e. the equivalent European price).

2 L. B. Andersen (1999). "A simple approach to the pricing of Bermudan swaptions in the multi-factor libor market model". In: *SSRN Electronic Journal*. DOI: 10.2139/ssrn.155208

Results

Summary: QLBS accurately prices exotic interest rate derivative products while providing efficient hedging on the underlying instruments.

1. Recovering zero coupon bond (ZCB) prices
2. Replicating forward swap values
3. Pricing European swaptions via Monte Carlo
4. Pricing and hedging via QLBS

Recovering ZCB Prices

To validate short-rate evolution we evaluate (6) which is the generic bond pricing formula.

$$P(0, T) = exp(-\int_0^T r(\tau)d\tau) \tag{6}$$

Rel. Error - QuantLib (bps)

	1	3	6	12	24	60	120
5	0.00	0.08	0.09	0.09	-0.48	-0.35	2.09
10	0.00	0.09	0.10	0.10	-0.46	-0.94	-1.75

Recovering Swap Prices

To validate swap pricing, we replicate vanilla instruments through QuantLib libraries and price under equivalent conditions.

We consider 5 year and 10 year swaps below with varying forward times.

Rel. Error - QuantLib (%)

	1	3	6	12	24	60	120
5	-0.19	0.15	-0.32	0.07	-0.3	0.087	-0.33
10	0.13	00.16	0.31	0.25	0.35	0.80	-0.5

Recovering Swaption Prices via Monte Carlo

We compare Monte Carlo derived European swaption prices with those derived analytically from Quantlib (which uses Jamishidan swaption pricing).

We consider options into 5 and 10 year swaps with varying expiry (chosen based on which instruments were most liquid).

Rel. Error - QuantLib (%)

	1	3	6	12	24	60	120
5	-6.78	-10.18	-5.66	-10.19	-5.19	-8.75	-2.27
10	-8.17	-3.04	-7.91	-4.08	-6.73	-4.13	-7.61

Pricing European and Bermudan Swaptions using QLBS

Rel. Error - Euro (%)

	1	3	6	12	24	60	120
5	-13.2	-12.9	-10.0	-0.6	-7.3	-2.0	3.6
10	-15.1	-15.3	-14.3	-10.3	-11.0	-5.9	N/A

Rel. Error - Berm (%)

	1	3	6	12	24	60	120
5	9.5	-6.3	-15.7	-14.1	-11.5	-6.7	-1.9
10	-1.65	-2.01	-15.69	-22.5	-20.8	-16.3	N/A

Future Directions

With additional time I would explore:

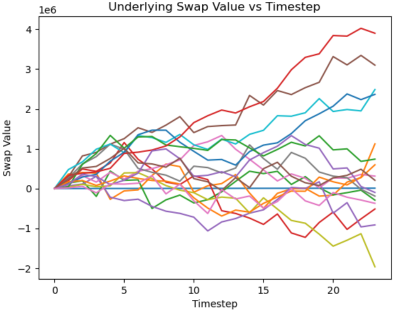
1. Multi-factor models for term structure evolution
2. Non-constant volatility models
3. Sensitivity of price/hedge to bumps in yield curve
4. Fitted Q-learning and model-free, off-policy RL
5. More exotic agreements - i.e. varying fixed/floating payment schedules, strike rates, etc.
6. Dynamic risk-free rates
7. Incorporating dividends into reward structure

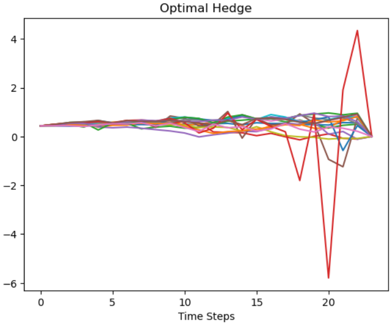
Acknowledgements

Thank you to Sandeep, Dongsheng, Ying, Yan, Yuncai, Dorothy, and the entire Strats team for their help and guidance!

Appendix 1.1

1M x 5y swap: Underlying Trajectories and Optimal Hedging

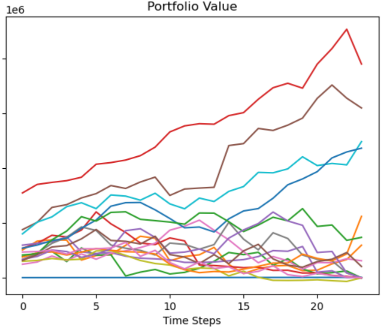


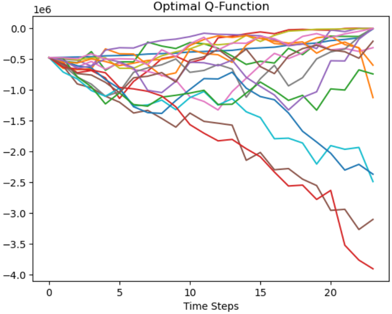


Swap value deflated by money-market account (risk-neutral numeraire) is a martingale.

Appendix 1.2

1M x 5y swap: Portfolio Values and Q-Values

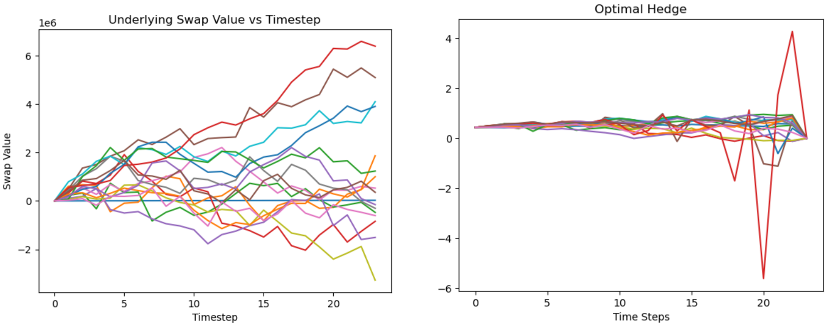




Observe expected increase in portfolio value and near-monotonic decrease in -Q value (increase +Q).

Appendix 1.3

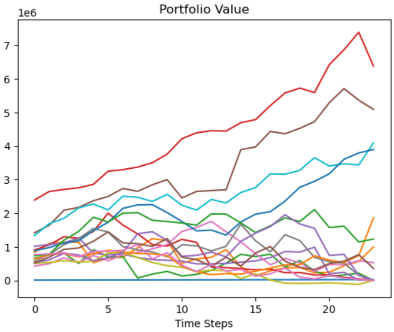
1M x 10y swap: Underlying Trajectories and Optimal Hedging

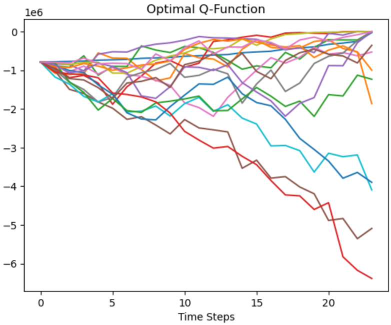


In real-world settings, transaction costs would enforce smooth hedging which can be achieved with certain regularization on optimal actions.

Appendix 1.4

1M x 10y swap: Portfolio Values and Q-Values





Appendix 2.1

Evaluating Hedge Effectiveness - 1M x 5y [3]

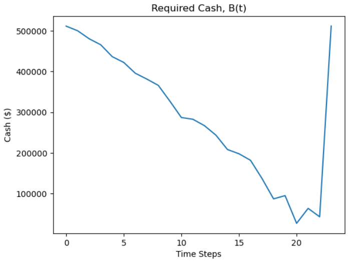


Figure: On notional of \$100M, we require an initial bank account of \$500K for optimal hedging.

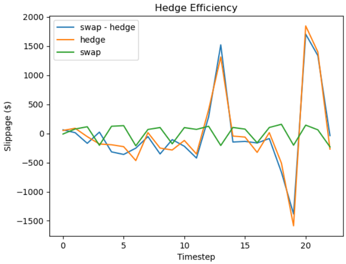


Figure: Our total slippage, i.e. unprotected ΔV_{swap} is \$9,822.61 or \$427.07 per timestep.

3 J. H. Hoencamp, S. Jain, and D. Kandhai (2022). "A semi-static replication approach to efficient hedging and pricing of callable IR derivatives". In

Appendix 2.2

Evaluating Hedge Effectiveness - 1M x 10y

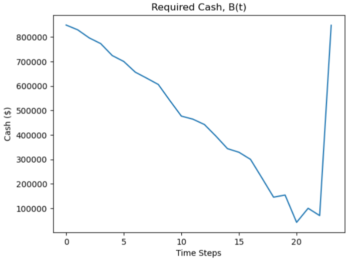


Figure: On notional of \$100M, we require an initial bank account of \$800K for optimal hedging.

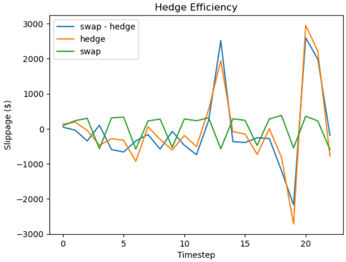


Figure: Our total slippage, i.e. unprotected ΔV_{swap} is $\$16,334.83$ or \$710.21 per timestep.

Appendix 2.3

Another look at Hedge Effectiveness

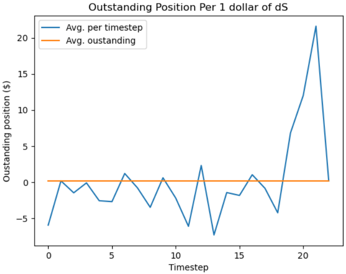


Figure: 1M x 5y outstanding per 1$ change in underlying (ΔS).

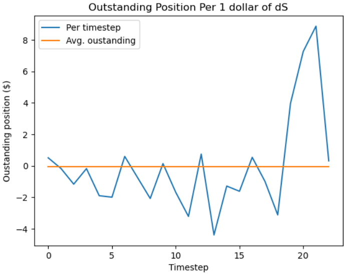


Figure: 1M x 10y outstanding per 1$ change in underlying (ΔS).

We assume risk per-timestep, however our average net-position per dollar change in the underlying is approx. $0.2.

Appendix 3

Sparse first-differences with L1 regularization. We add weighted L1 regularization to a_t^*, and vectorize w.r.t. time.

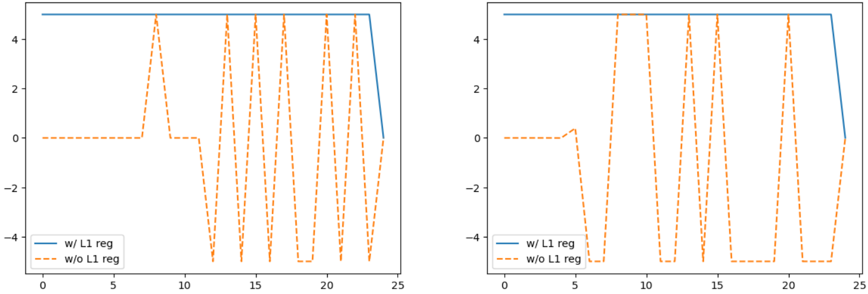


Figure: L1 regularization yields minimal optimality gap with sparse hedge first differences; shown for coefficients related to two features.

Appendix 4

Implementation details. Some parameters and assumptions were made to normalize experimental results.

- Fixed and floating-leg payment schedules identical
- Quarterly payments
- Hedge re-balancing across 24 timesteps
- 8192 Sobol (low-discrepancy) sequences were used to simulate short-rate paths (computation time around 30 secs)
- 3% risk-free rate
- All swaptions were evaluated ATM
- Purely risk-based hedging

Bibliography

Andersen, L. B. (1999). "A simple approach to the pricing of Bermudan swaptions in the multi-factor libor market model". In: *SSRN Electronic Journal*. DOI: 10.2139/ssrn.155208.

Halperin, I. (2019). *QLBS: Q-Learner in the Black-Scholes(-Merton) Worlds*. arXiv: 1712.04609 [q-fin.CP].

Hoencamp, J. H., S. Jain, and D. Kandhai (2022). "A semi-static replication approach to efficient hedging and pricing of callable IR derivatives". In.