Market Microstructure & Optimized Trade Execution using RL 31-Oct-20

G. Belford

Abstract

Based on GEN/ETH market microstructure LOB data, trades, objectives & constraints – recommend suitable algos which maximize pnl/minimize inventory losses within a fixed time horizon. The Reinforcement Learning (RL) method is chosen to help achieve optimized trade execution relative to given constraints. This ML method searches for signals to help predict short-term price direction (up, neutral, down), execution likelihood at a given price, buy/sell order imbalance, etc., all of which are used to reprice our orders. RL algo results are compared vs. those of a standard market-making algo bot.

Objectives

A market maker's economic objective is to maximize the frequency of crossing trades, completing the transaction cycle of buying/selling assets in order to capture the bid-ask spread.

Specific Objectives

Objective1. Maximize time present (ie. quoting - aim for constant quoting) with a minimum of 2 orders on the book – 1 on each side (no constraint on how far from mid). Note constraint1 below.

Objective2. Always 1 open order at least on 1 side of the book.

Objective3. Inventory – maintain balance as far as possible & minimize losses over period (3 weeks). This will be implemented via inclusion of real-time pnl mgmt, constraint2 & inventory skew – all detailed below.

Fig. GENETH real-time pnl mgmt (in USD fiat).

Fig. Inventory PnL Risk Mgmt		GEN	GEN/ETH Traded Px & Remaining B/E Px	GENETH Current Mkt Px	ETH Position at Current Market Px if unwind trade		ETH	ETH	
Trade1. MM Buy GEN/ETH spread	GEN long position	50,000	0.00038001			ETH short position	(19)	Current Market Offer	400
Trade2. MM Sell GEN/ETH spread	GEN short position	(25,000)	0.00039499			ETH long position	10	Current Market Bid	390
	Net GEN position (long)	25,000	0.00036503	0.00040000	10	Net ETH Residual Position from 2 trades	(9)	After flatten GEN/ETH spread then need to buy this ETH residual position back to close into USD fiat.	

USD fiat open	0	(a)	As we are Net Long GEN/ETH & have +MtM, if we flatten GEN/ETH position then we also need to sell our residual ETH at ETH market bid price to convert pnl back to USD	
USD PnL on residual ETH closeout	341	(b)		1
USD fiat close	341	(a) + (b)		

Constraints

Constraint1. Maintain an average spread (over review period) of 12% on market (nb. GENETH LOB average BBO spread is 23%). A sample average spread calc is shown below (our order postings are highlighted in yellow, resulting in an average spread of 8%). Aggregate LOB data stats are also displayed.

Orderbook

$$Spread\ (in\ \%) = 2 * \frac{(Max(Sell) - Min(Buy))}{Max(Sell) + Min(Buy)} x\ 100$$

timestamp	order_id	bid3_vol	bid2_vol	bid1_vol	bid3_px	bid2_px	bid1_px	ask1_px	ask2_px	ask3_px	ask1_vol	ask2_vol	ask3_vol	MM Spread (in %)
2019/08/01 00:	9e63fe0	1,200	41	86	0.00031089	0.00035582	0.00037559	0.00038132	0.00038133	0.00038324	1,028	1,983	1,923	7%
2019/08/01 00:	9e63fe0	1,200	44	41	0.00031089	0.00035582	0.00037559	0.00038132	0.00038133	0.00038324	1,028	1,983	1,923	2%
2019/08/01 00:	9e63fea	41	88	41	0.00031089	0.00035582	0.00037559	0.00038132	0.00038133	0.00038324	1,028	1,983	1,923	2%
2019/08/01 00:	9e63fee	41	44	86	0.00031089	0.00035582	0.00037559	0.00038132	0.00038133	0.00038324	1,028	1,983	1,923	21%
													Average:	8%

Fig. Aggregate LOB Data for top 5 levels each side	Max	Average	Min
bidvolume_total	148,326	28,932	144
askvolume_total	25,572	6,803	211
bidoverask_volume_total_ratio	310	6	0.020
ask4_5_spread%	125%	15%	0%
ask3_4_spread%	165%	14%	0%
ask2_3_spread%	165%	10%	0%
ask1_2_spread%	165%	7%	0%
bid1_ask1_spread_abs	0.000420	0.000066	0.000000
bid1_ask1_spread_%	**189%**	**23%**	**0%**
midspread_price	0.000497	0.000327	0.000224
bid1_2_spread%	45%	3%	0%
bid2_3_spread%	44%	5%	0%
bid3_4_spread%	3095%	12%	0%
bid4_5_spread%	3095%	257%	0%

Constraint2. Max Inventory "GEN/USD" & “ETH/USD" (each USD 10k equiv).

Trading the GEN/ETH spread does not impact the USD fiat balance. To move to USD fiat the spread position is unwound by simply offsetting the # of contracts traded (ex. buy then sell 100 GEN/ETH) & close out the residual ETH/USD position into USD. Min/Max USD 10,000 inventory position limits (Fig below) are monitored. Once one side is hit the bot will stop quoting that side of the market.

Inventory: GEN/USD						**GEN**		
GEN Long	GEN Short	GEN Net	GEN/USD Mkt Bid	GEN/USD Mkt Ask	GEN/USD Mid Px	**USD equiv. Exposure Limit: USD 10,000**	Trade_ID	Time_Stamp
50,000		50,000	0.0705127	0.0705827	0.0705477	3,527	382652400_1	2019/08/01 22:19:56:1956
	25,000	25,000	0.0706023	0.0706100	0.0706062	1,765	382652500_1	2019/08/01 22:19:59:1957

Inventory: ETH/USD						**ETH**		
ETH Long	ETH Short	ETH Net	ETH/USD Mkt Bid	ETH/USD Mkt Ask	ETH/USD Mid Px	**USD equiv. Exposure Limit: USD 10,000**	Trade_ID	Time_Stamp
	(19)	(19)	178.69	185.55	182.12	(3,460)	382652400_2	2019/08/01 22:19:56:1956
10		(9)	177.69	184.55	181.12	(1,653)	382652500_2	2019/08/01 22:19:59:1957

Inventory Skew

Marketmakers current inventory & target level result in bias for taking more or less risk. This bias is expressed via the marketmaker's quote, order size, or both:

Scen1. Inventory_current > Inventory_target ->

spread_bid > spread_ask, ordersize_bid < ordersize_ask

Scen2. Inventory_current < Inventory_target ->

spread_bid < spread_ask, ordersize_bid > ordersize_ask

P_bid = P_ref – bid_spread where bid_spread is spread (Ref Px-> Bid Px)

P_ask = P_ref + ask_spread where ask_spread is spread (Ref Px->Ask Px)

bid-ask_spread = P_ask – P_bid = spread_bid + spread_ask

A risk adverse market maker (Scen1) where current Inventory exceeds target will skew the params of its orders to decrease the probability and/or magnitude of acquiring more inventory relative to selling inventory.

RL algo – concept

A review of public research into market microstructure identifies RL as a useful tool for optimized trade execution. Our study involves training an agent on historical GEN/ETH LOB and trade data in order for it to take actions maximizing the reward function (pnl). This agent is fed many "states" and has the goal of profit optimization. A simulator (combining historical real order flow & artificial orders generated by the RL algo) will execute orders, maintain order book priority and monitor execution costs/uncertainties (ie. bid-ask spread, market impact, non-execution risks). Both private and market variables are considered. 3 weeks of LOB data & trade data is split into a training set (2 weeks) & a test set (1 week).

RL algo – implementation

Training

Keeping Objective1 (constant quoting) in mind & the fact we have millisecond LOB data, we will feed the RL 2 parameters - order sizes ("Orders", say 30, 50 & 100 contracts) and execution horizon ("Execution", say 30s, 1min & 2 min). The RL algo will iterate over every combo of these 2 parameters for every state space representation, taking into account 2 private variables (remaining amount of time/decision points t/T & inventory i/I in the episode). During *training* our policy actions (our trades) will have no impact on subsequent LOB evolution. Therefore for a given order book state, our trade simulation results – executed trades – are used to update our inventory, but the order book itself (market variable) will evolve without our trading results. During *testing* this assumption is removed.

We will then add market variables to our state space to investigate whether optimal actions should be contingent on market conditions. Bid-ask volume misbalance will be used in this regard (GENETH LOB stats below).

Fig. Aggregate Top 5 Bid Volume > Aggregate Top 5 Ask Volume	
% Time Bid Volume > Ask Volume	86%
% Time Bid Volume < Ask Volume	14%

Testing

RL-optimized execution policy performance will then be empirically compared (using 1 week test set) against a baseline execution strategy. A standard market-making bot - Github BitMEX XBTUSD MM Bot – will be modified to perform baseline analysis examining inventory pnl, trading costs (in bp over mid spread price).

Appendix A. LOB stats

raw_orderbooks_geneth.json
1. Exchange: Ethfinex
2. LOB Pair: GEN/ETH (DAOStack/Ethereum)
3. HistData: 3 weeks (2019/08/01 00:00:01:01 - 2019/08/20 23:59:58:5958 (293369 rows) UNIX (13 digit) millisecond timescale data.
4. Bid-Ask Spread (BBO) average: .000066
5. Bid-Ask Spread (BBO) average: 23.17% but can spike much higher (max 189%).
6. Bid Depth: 22 levels - price & volume quote frequency spotty below 10 levels.
7. Ask Depth: 25 levels - consistently showing levels.
8. Overall Ask & Bid volume stats shown below.
9. Overall Bid volumes skewed with some large posted volumes at 5th level & below.

	Bid	Ask
Max	409k	32k
Average	334k	15k
Min	118k	7k

Appendix B. Trade stats

ethfinex_geneth_publictrades.csv
Description: Public trades done on the order book over the LOB json review period.
1. Originally 764 trades in raw file.
2. Trades ordered sequentially by trade_order_id as some cases where trades done at exactly same timestamp.
3. Ordering by trade_order_id confirmed by spot checks.
Examples eyeballed make sense as can see order_book posts/volume level shifts changing as expected.
4. 59 trades removed at the end of file as were timestamped > LOB json final timestamp (2019/08/20 23:59:58:5958).
5. 1 trade (trade_order_id: 385700236) Buy 0 @ 0.000315 - removed.
6. Similarly, 3 sells (trade_order_ids: 382652463, 383116223, 383299678) with 0 volume removed as well.
7. 701 trades thus actually used: # Buys: 320 (Total Volume: 136,196), # Sells: 381 (Total Volume: 178,212)
8. Trade direction (buy/sell) shown from perspective of trader - not market maker, so for example a (trader) sell may be seen to be done at the market maker bid.
9. Average Volume at low end of range (Buys: 426, Sells: 503)

Trade stats	Buys	Sells
Trade_Count (final-used)	320	381
Max Volume per trade	5,804	6,545
Average Volume	426	503
Min Volume	1	1
Total Volume traded	136,196	178,212
Total Traded Notional	41.204	54.687

Appendix C. RL notes (off internet)

Q learning is a subset of RL where you look at the probdist of responses to various actions. This kind of ML is distinct from statistics. RL comes from optimization & decision science. So the best ML practitioners in finance may not be statisticians at all. RL is best suited to financial markets. In

Supervised Learning you don't account for the fact that your decision changes the state of the world. You are observing the data and making decisions – nothing about your decision feeds back into the market. This isn't the case in the markets, so by definition RL seems most appropriate.