Ekeoma Osondu

Blockhouse QR Work Trial Reflection

Problem 1: Modeling Choices

The temporary-impact function $g_t(x)$ is the per-share slippage one suffers if, at time t, you submit a market order for x shares and must walk up the limit order book to fill it.

Let the midprice be $m_t = \frac{\text{ask\_px}_{t,0}+\text{bid\_px}_{t,0}}{2}$; for a buy, you consume the ask ladder level by level, accumulating size until the cumulative quantity first reaches x. I will denote the ask prices and sizes by $\{(p^{\text{ask}}_{t,j}, v^{\text{ask}}_{t,j})\}_{j=0}^{J-1}$ and the cumulative depth by $Q_{t,k} = \sum_{j=0}^{k-1} v^{\text{ask}}_{t,j}$ with $Q_{t,0} = 0$.

The dollar cost to execute x shares is the sum of full levels plus the partial fill on the last level, and the per-share impact is

$$g_t(x) \;=\; \frac{1}{x}\left[\sum_{k=1}^{\ell-1} p^{\text{ask}}_{t,k-1}\big(Q_{t,k} - Q_{t,k-1}\big) \;+\; p^{\text{ask}}_{t,\ell-1}\big(x - Q_{t,\ell-1}\big)\right] - m_t, \qquad Q_{t,\ell-1} < x \le Q_{t,\ell}.$$

Since deeper levels typically quote worse prices, $g_t(x)$ is piecewise linear, with kinks at the cumulative-depth breakpoints $\{Q_{t,k}\}$. A single global linear approximation $g_t(x) \approx \beta_t x$ is way more algebraically convenient but can be inaccurate for large x: the first shares primarily pay the spread while later shares face sharply rising marginal prices.

Instead a convex piecewise-linear is better for encoding in a linear program via segment variables and extract a local slope only when a number is required. Estimation is simple in this context, at each time bucket, read the top-J levels, compute $\{Q_{t,k}\}$ and segment slopes $s_{t,k} = p^{\text{ask}}_{t,k-1} - m_t$, and then fully characterize $g_t(\cdot)$.

Problem 2: Choosing my Trade Schedule $\{x_i\}_{i=1}^N$

We split the trading day into $N = 390$ one–minute buckets and buy a total of S shares:

$$\sum_{i=1}^{N} x_i = S, \qquad x_i \ge 0.$$

Let $C_i(x)$ be the *dollar* temporary–impact cost of executing x shares in bucket i. Our goal is

$$\min_{\{x_i\}_{i=1}^N} \sum_{i=1}^{N} C_i(x_i) \quad \text{s.t.} \quad \sum_{i=1}^{N} x_i = S,\; x_i \ge 0.$$

(a) Piecewise–Linear Cost to Linear Program

From the order book at time i, build a convex PWL cost curve by splitting depth into segments:

$$0 = Q_{i,0} < Q_{i,1} < \cdots < Q_{i,K_i}, \qquad s_{i,k} = \text{marginal slippage on segment } k.$$

If $z_{i,k}$ is the number of shares executed on segment k, then

$$C_i(x_i) = \sum_{k=1}^{K_i} s_{i,k} z_{i,k}, \qquad x_i = \sum_{k=1}^{K_i} z_{i,k}, \qquad 0 \le z_{i,k} \le Q_{i,k} - Q_{i,k-1}.$$

Stacking all minutes gives us the linear program:

$$\begin{aligned} \min_{\{z_{i,k}\}} \quad & \sum_{i=1}^{N} \sum_{k=1}^{K_i} s_{i,k}\, z_{i,k} \\ \text{s.t.} \quad & \sum_{i=1}^{N} \sum_{k=1}^{K_i} z_{i,k} = S, \\ & 0 \le z_{i,k} \le Q_{i,k} - Q_{i,k-1} \quad \forall\, i, k. \end{aligned}$$

After solving, recover $x_i = \sum_k z_{i,k}$. This LP is a few thousand variables at most which isnt too expensive to solve.

(b) Linearized Cost to Quadratic Program

If we use a local linear model $g_i(x) \approx \beta_i x$ the dollar cost is

$$C_i(x) = x \cdot g_i(x) \approx \beta_i x^2,$$

so the problem becomes a convex quadratic program:

$$\min_{x_i \ge 0} \sum_{i=1}^{N} \beta_i x_i^2 \quad \text{s.t.} \quad \sum_{i=1}^{N} x_i = S.$$

Once again, it is not too large and can be solved using common quadratic programming techniques (Lagrange Multipliers etc).