

Cryptocurrency Research Platform

1. Data Pipeline

Price Matrix

Raw OHLCV data is fetched and processed into a price matrix:

$$P_{t,i} \in \mathbb{R}^{T \times N}$$

where $t \in [1, T]$ are timestamps and $i \in [1, N]$ are cryptocurrency symbols.

Returns Computation

Simple returns are computed as:

$$r_{t,i} = \frac{P_{t,i} - P_{t-1,i}}{P_{t-1,i}}$$

2. Signal Strategies

2.1 Momentum Strategy

Captures trending behavior using past returns:

$$\text{Signal}_{\text{momentum},t,i} = \frac{P_{t-1,i} - P_{t-L-1,i}}{P_{t-L-1,i}}$$

where L is the lookback period (default: 168 hours = 1 week).

2.2 Mean Reversion Strategy

Exploits price extremes using z-score of returns:

$$\mu_{t,i} = \frac{1}{L} \sum_{k=t-L}^{t-1} r_{k,i}$$

$$\sigma_{t,i} = \sqrt{\frac{1}{L} \sum_{k=t-L}^{t-1} (r_{k,i} - \mu_{t,i})^2}$$

$$z_{t,i} = \frac{r_{t-1,i} - \mu_{t-1,i}}{\sigma_{t-1,i}}$$

$$\text{Signal}_{\text{mean_rev},t,i} = -z_{t,i}$$

The negative sign converts oversold conditions (negative z-score) into positive signals for mean reversion.

2.3 EWMA Crossover Strategy

Identifies trend changes via exponential moving average divergences:

$$\text{EWMA}_{\text{fast},t,i} = \alpha_f \cdot P_{t,i} + (1 - \alpha_f) \cdot \text{EWMA}_{\text{fast},t-1,i}$$

$$\text{EWMA}_{\text{slow},t,i} = \alpha_s \cdot P_{t,i} + (1 - \alpha_s) \cdot \text{EWMA}_{\text{slow},t-1,i}$$

where $\alpha_f = \frac{2}{\text{fast_window}+1}$ and $\alpha_s = \frac{2}{\text{slow_window}+1}$.

Rolling standard deviation:

$$\sigma_{\text{rolling},t,i} = \sqrt{\frac{1}{W} \sum_{k=t-W}^{t-1} (P_{k,i} - \bar{P}_{t,i})^2}$$

Final signal (normalized by volatility):

$$\text{Signal}_{\text{EWMA},t,i} = \frac{\text{EWMA}_{\text{fast},t-1,i} - \text{EWMA}_{\text{slow},t-1,i}}{\sigma_{\text{rolling},t-1,i}}$$

3. Predictor (Cross-Sectional Processing)

For each strategy s , the predictor transforms raw signals into standardized signals.

3.1 Cross-Sectional Z-Score

Normalize signals across all assets at each timestamp:

$$\bar{S}_{s,t} = \frac{1}{N} \sum_{i=1}^N \text{Signal}_{s,t,i}$$

$$\sigma_{s,t} = \sqrt{\frac{1}{N} \sum_{i=1}^N (\text{Signal}_{s,t,i} - \bar{S}_{s,t})^2}$$

$$Z_{s,t,i} = \frac{\text{Signal}_{s,t,i} - \bar{S}_{s,t}}{\sigma_{s,t}}$$

This produces the **unfiltered signal** (z-scored but not filtered).

3.2 Quantile Filtering

Filter assets based on cross-sectional quantiles:

$$Q_{\text{top}}(t) = \text{quantile}(Z_{s,t,:}, q_{\text{top}})$$

$$Q_{\text{bottom}}(t) = \text{quantile}(Z_{s,t,:}, q_{\text{bottom}})$$

$$F_{s,t,i} = \begin{cases} Z_{s,t,i} & \text{if } Z_{s,t,i} \geq Q_{\text{top}}(t) \text{ (long)} \\ Z_{s,t,i} & \text{if } Z_{s,t,i} \leq Q_{\text{bottom}}(t) \text{ (short)} \\ 0 & \text{otherwise} \end{cases}$$

Default: $q_{\text{top}} = 0.8$ (top 20%), $q_{\text{bottom}} = 0.2$ (bottom 20%).

3.3 Re-Zscore on Active Names

Re-normalize only the non-zero (active) positions:

$$\mathcal{A}_t = \{i : F_{s,t,i} \neq 0\}$$

$$\bar{F}_{s,t} = \frac{1}{|\mathcal{A}_t|} \sum_{i \in \mathcal{A}_t} F_{s,t,i}$$

$$\sigma_{F,t} = \sqrt{\frac{1}{|\mathcal{A}_t|} \sum_{i \in \mathcal{A}_t} (F_{s,t,i} - \bar{F}_{s,t})^2}$$

$$\tilde{Z}_{s,t,i} = \begin{cases} \frac{F_{s,t,i} - \bar{F}_{s,t}}{\sigma_{F,t}} & \text{if } i \in \mathcal{A}_t \\ 0 & \text{otherwise} \end{cases}$$

4. Portfolio Construction

4.1 Combine Unfiltered Signals

Combine z-scored (unfiltered) signals from multiple strategies:

$$C_t^{\text{unfiltered}} = \sum_{s=1}^S w_s \cdot Z_{s,t}$$

where w_s are strategy allocation weights with $\sum_{s=1}^S w_s = 1$.

4.2 Filter Combined Signal

Apply quantile filtering to the combined signal:

$$\begin{aligned} Q_{\text{combined, top}}(t) &= \text{quantile}(C_t^{\text{unfiltered}}, q_{\text{combined, top}}) \\ Q_{\text{combined, bottom}}(t) &= \text{quantile}(C_t^{\text{unfiltered}}, q_{\text{combined, bottom}}) \\ C_t^{\text{filtered}} &= \begin{cases} C_t^{\text{unfiltered}} & \text{if } C_t^{\text{unfiltered}} \geq Q_{\text{combined, top}}(t) \\ C_t^{\text{unfiltered}} & \text{if } C_t^{\text{unfiltered}} \leq Q_{\text{combined, bottom}}(t) \\ 0 & \text{otherwise} \end{cases} \end{aligned}$$

4.3 Re-Zscore and Generate Weights

Re-zscore the combined filtered signal on active names to produce final portfolio weights:

$$w_{t,i} = \tilde{C}_{t,i}$$

where $\tilde{C}_{t,i}$ is computed using the re-zscore procedure from Section 3.3.

5. Backtesting

5.1 Timing Convention

- $w_{t,i}$ = portfolio weights determined at close of day t
- $r_{t+1,i}$ = return from close of day t to close of day $t + 1$
- Portfolio return: $R_{t+1} = \sum_{i=1}^N w_{t,i} \cdot r_{t+1,i}$

5.2 Rebalancing Schedule

Weights are held constant between rebalancing periods:

$$w_{t,i}^{\text{rebalanced}} = \begin{cases} w_{t,i} & \text{if } t \bmod f_{\text{rebal}} = 0 \\ w_{t-1,i}^{\text{rebalanced}} & \text{otherwise} \end{cases}$$

where f_{rebal} is the rebalancing frequency (e.g., 24 hours for daily).

5.3 Transaction Costs

Turnover at time t :

$$\text{Turnover}_t = \frac{1}{2} \sum_{i=1}^N |w_{t,i}^{\text{rebalanced}} - w_{t-1,i}^{\text{rebalanced}}|$$

Transaction costs:

$$\text{TC}_t = \text{Turnover}_t \times (\text{cost}_{\text{bps}} + \text{slippage}_{\text{bps}}) \times 10^{-4}$$

Net portfolio return:

$$R_t^{\text{net}} = R_t - \text{TC}_t$$

5.4 Cumulative Returns

$$\text{CumRet}_t = \prod_{k=1}^t (1 + R_k^{\text{net}})$$

6. Performance Metrics

6.1 Annualized Return

Using geometric mean:

$$R_{\text{annual}} = (1 + R_{\text{total}})^{\frac{P}{T}} - 1$$

where $R_{\text{total}} = \prod_{t=1}^T (1 + R_t^{\text{net}}) - 1$ and P is periods per year (8760 for hourly data).

6.2 Annualized Volatility

$$\sigma_{\text{annual}} = \sigma(R^{\text{net}}) \times \sqrt{P}$$

where $\sigma(R^{\text{net}})$ is the standard deviation of net returns.

6.3 Sharpe Ratio

$$\text{Sharpe} = \frac{R_{\text{annual}} - r_f}{\sigma_{\text{annual}}}$$

where r_f is the risk-free rate (default: 0).

6.4 Maximum Drawdown

$$\text{Peak}_t = \max_{k \leq t} \text{CumRet}_k$$

$$\text{DD}_t = \frac{\text{CumRet}_t - \text{Peak}_t}{\text{Peak}_t}$$

$$\text{MaxDD} = \min_t \text{DD}_t$$

6.5 Calmar Ratio

$$\text{Calmar} = \frac{R_{\text{annual}}}{|\text{MaxDD}|}$$

6.6 Win Rate

$$\text{WinRate} = \frac{\#\{t : R_t^{\text{net}} > 0\}}{T}$$

6.7 Profit Factor

$$\text{ProfitFactor} = \frac{\sum_{t: R_t^{\text{net}} > 0} R_t^{\text{net}}}{\left| \sum_{t: R_t^{\text{net}} < 0} R_t^{\text{net}} \right|}$$

7. Target Engineering

Forward Returns (for evaluation)

$$\text{Target}_{t,i} = r_{t+h,i} = \frac{P_{t+h,i} - P_{t,i}}{P_{t,i}}$$

where h is the forward period (default: 1 hour).

Pipeline Summary

The complete workflow is:

1. **Data** → Price matrix $P_{t,i}$
2. **Signals** → Raw signals for each strategy s
3. **Predictor** → Raw signal → Z-score → Unfiltered signal $Z_{s,t,i}$
4. **Portfolio** → Combine unfiltered signals → Filter → Re-zscore → Weights $w_{t,i}$
5. **Backtest** → Apply weights with rebalancing → Compute returns with costs → R_t^{net}
6. **Metrics** → Evaluate performance (Sharpe, MaxDD, etc.)