Maroon Capital Winter 2022 Trading Project

Team Lead: Frank Li

Jerry Liu, Hongli Zhao

March 7, 2022

Overview

- Two types of strategies:
  1. Trend-following: a trend will persist. Buy in an uptrend, sell in a downtrend (e.g. moving average)
  2. Contrarian / mean-reverting: trend tops out and reverses. Buy in a downtrend, sell in an uptrend (e.g. oscillators)
- Assumption: stock prices are (to some degree) predictable; EMH is incorrect
- Most important task is to capture trend: choose stock with high seasonality and low volatility

Choosing Stocks

- We focus on large-cap companies in the retail business, which include food retail (e.g. WalMart), department stores (e.g. Costco), drugstores (e.g. CVS), discount stores (e.g. Dollar Tree).
- We train using roughly 8 years of data (01/03/2012 – 12/31/2021) and use a 80-20 train-test split. All reported performance are on test data
- To check for high seasonality, we consider the adjusted close price as a time series X_t and perform the STL decomposition [R.B. Cleveland] (1990):

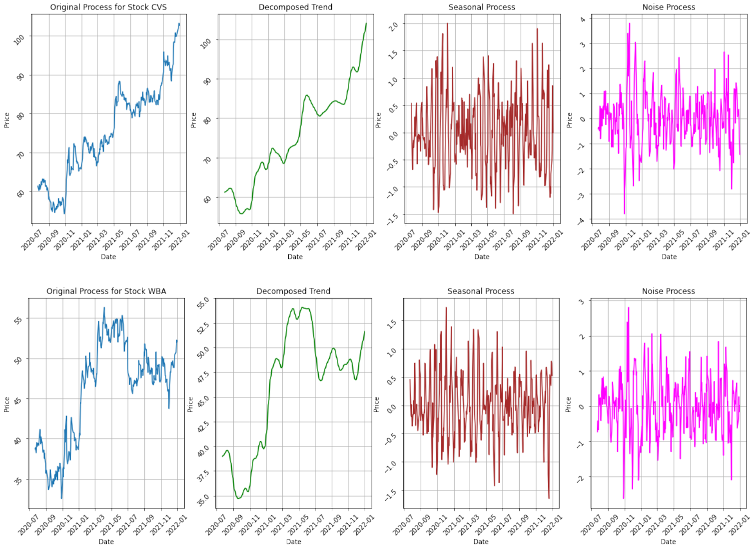
$$X_t = \underbrace{S_t}_{\text{seasonal component}} + \underbrace{T_t}_{\text{trend component}} + \underbrace{R_t}_{\text{noise}}$$

  then we measure strength of seasonality as:

$$s := \max\left\{0, 1 - \frac{\mathbb{V}ar\left[R_t\right]}{\mathbb{V}ar\left[S_t + R_t\right]}\right\}$$

  where variance is estimated from data. By definition $s \in [0, 1]$. The closer s to 1, the stronger the seasonality.
- The volatility will simply be standard deviation.

- We rank the 27 stocks based on these two criterion and favor 3 with highest seasonality and lowest volatility



- Based on inspection, we choose:
-

Name	Seasonality	Volatility
WBA	0.7698	14.5249
BBWI	0.7693	14.9787
KR	0.7605	8.5052

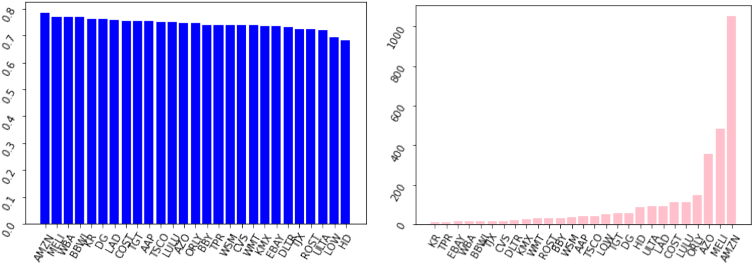


Figure: Left: Seasonality; Right: Volatility over 2517 date points

1. Walgreens Boots Alliance (WBA): owns the second largest pharmacy store chain Walgreens.
2. Bath & Body Works Inc (BBWI): owns the retail store chain specializing in soap, lotion, fragrances and candles
3. Kroger Co. (KR) is an NYSE-listed retail company operating food processing or manufacturing facilities, supermarket fuel centers, pharmacies, in-store medical clinics. The largest supermarket chain by revenue

Trading Strategies

- We survey a few common strategies:

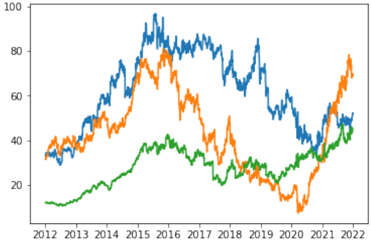


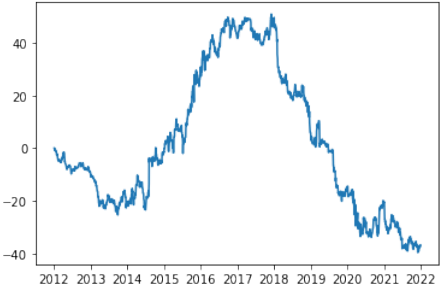
Figure: WBA, BBWI, KR

  - Bollinger Band Breakout
  - Moving Average Crossover
  - RSI Oscillator
  - MACD Crossover
- In the end, we also present Gaussian smoothing and polynomial regression as a way to isolate trend from noise, and leverage its smoothness to implement trend following and contrarian beliefs.
- In optimizing the parameters, we use grid search with reasonable ranges, seeking to maximize Sharpe ratio

Random Strategy

Buy, sell, or hold whenever I feel like it (randomly generate from $\{-1, 0, 1\}$).

- Test on WBA:

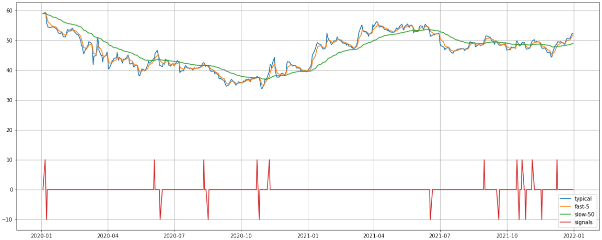


Moving Average Crossover

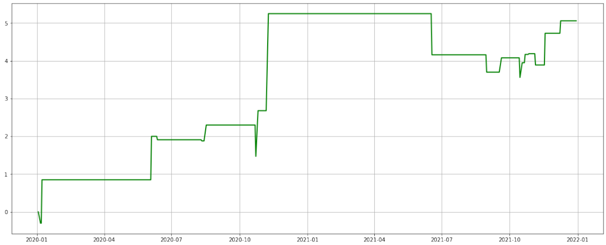
Use two different degrees of exponentially-weighted moving averages to smooth the price data (a "fast line" and a "slow line"). When fast line crosses the slow line upwards or downwards, generates signal.

```
if fast_line >= slow_line and not previously:
    generate buy
else if fast_line <= slow_line and not previously:
    generate sell
```

- Tested on WBA: grid search over $5, 10, 25$ for fast line; $50, 100, 200$ for slow line.



- Cumulative P&L:



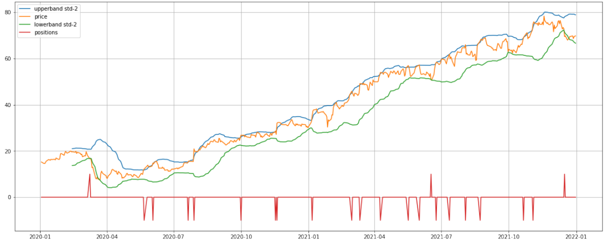
- Final Sharpe: 0.3081

Bollinger Band Breakout

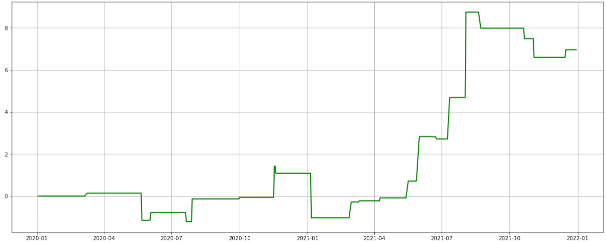
Wrap possible price fluctuations within a confidence interval ($\pm n\sigma$ standard deviations). If price broke out of the lower boundary, buy cheap; if price broke above the upper boundary, sell high.

```
if current_price <= lower_band and not previously:
    generate buy
else if current_price >= upper_band and not previously:
    generate sell
```

- Test on BBWI: grid search over $[5, 15, 20, 25, 30, 35, 40, 45, 50]$ for lookback period and standard deviation of $[1, 1.5, 2]$.
- Final parameters: 30 days lookback, 2 standard deviations



- Cumulative P&L:



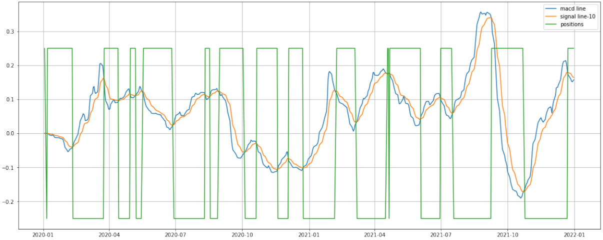
- Final Sharpe: 0.08892

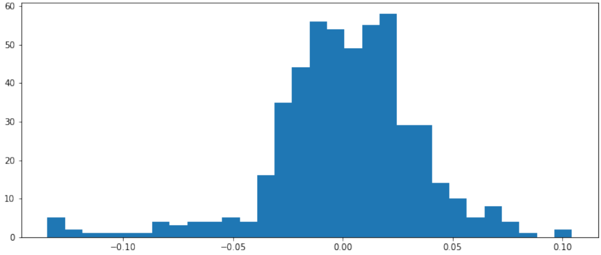
MACD Crossover

Measure divergence of fast line and slow line, then compute the signal line. MACD is a measure of momentum and the signal line is a lagged version; the crossing of the signal line by the MACD line indicates start of new trends.

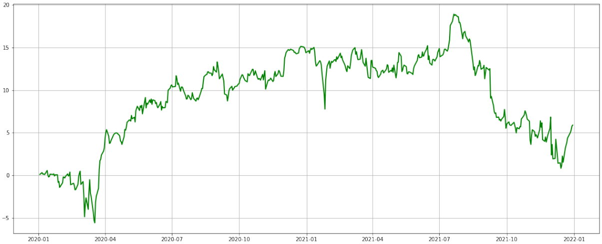
```
if MACD > signal:
    generate buy
else:
    generate sell
```

- Tested on KR. Parameters and grid search range:
  - fast line lookback: $[13, 16, 18, 20, 24, 26]$
  - slow line lookback: $[24, 26, 28, 30]$; make sure slow $\geq$ fast.
  - smoothing signal: $[8, 9, 10]$
- Final parameters: 26 days fast, 30 days slow, 10 days smoothing





- Cumulative P&L:



- Final Sharpe: 0.2836

RSI Osicllator

Relative Strength index is a momentum indicator comparing total strength of bullish and bearish movements.

$$RSI = 100 - \frac{100}{1 + \frac{\text{smoothed average gain over } x \text{ days}}{\text{smoothed average loss over } x \text{ days}}}$$

The smoothed gain / loss are computed as:

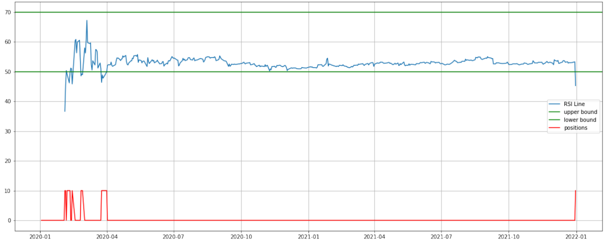
$$(\text{average over } x \text{ days}) \times (x - 1) + \text{last day's gain / loss}$$

Consider oversold if RSI above 70 and overbought if below 30; we decide to tighten the bounds.

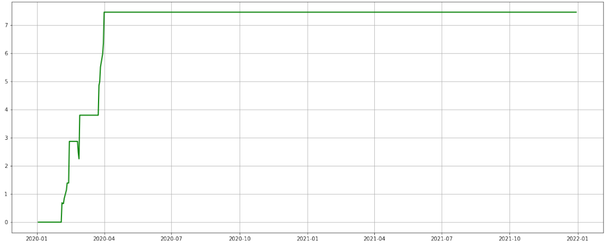
- Parameters:
  - interval: $[lb, ub]$ where $lb = [30, 40, 50, 55]$, $ub = [60, 65, 70]$.
  - lookback period: $[9, 14, 21]$

```
if RSI <= lb:
    generate buy
else if RSI >= ub:
    generate sell
```

- Tested on KR
  - Caveat: determining the interval is largely empirical; for too wide intervals, may not generate any signal at all
- Final parameters: lower bound = 50, upper bound = 70, lookback = 21.



- Cumulative P&L:



- Final Sharpe: 0.57

Smoothing

- A recurring theme is identifying the trend; loosely define "trend" as a smooth function that captures qualitative variations independent of noise.
  - Assumes:
  $$Y_t = f(X_t) + \epsilon_t$$
  where ϵ_t is noise. f is the function that we want to identify from empirical data of Y_t.
- If the trend were identified and is sufficiently smooth, we can consider the following momentum interpretations:
  1. (trend following) If $\frac{df}{dt} > 0$, we have an up trend, and we should buy; similarly for selling.
  2. (contrarian) If $\frac{df}{dt} = 0$, but $\frac{d^2f}{dt^2} > 0$, the stock has hit a bottom, and we should buy; similarly for selling
- We consider two ways of smoothing the data and test the two strategies.
  - Gaussian filtering
  - Polynomial regression

Gaussian Smoothing

- Given observations $y = (y_0, \cdots, y_N)$, smoothing relies on applying a filter $\mathcal{F}$:

$$\hat{y}(t) = \mathcal{F}(y(t)) = \sum_{i=0}^{N} w_i \cdot y(t-i)$$

  here w_i is a *kernel* which can also be a function of y, assigning different weights to observations of y.
- Example (simple moving average):

$$w_i = \frac{1}{n}\{i < n\}$$

  where n is the lookback period.
- Gaussian weights (unnormalized):

$$w(i, y) = \exp\left(-\frac{(y - y_i)^2}{2b^2}\right)$$

  where b is the bandwidth. The Gaussian function is a very smooth function that decreases from the mean.

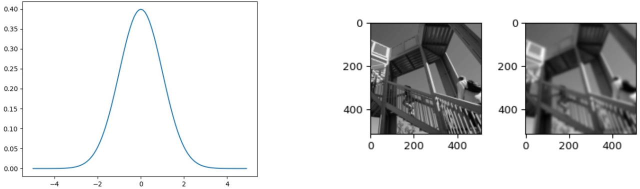
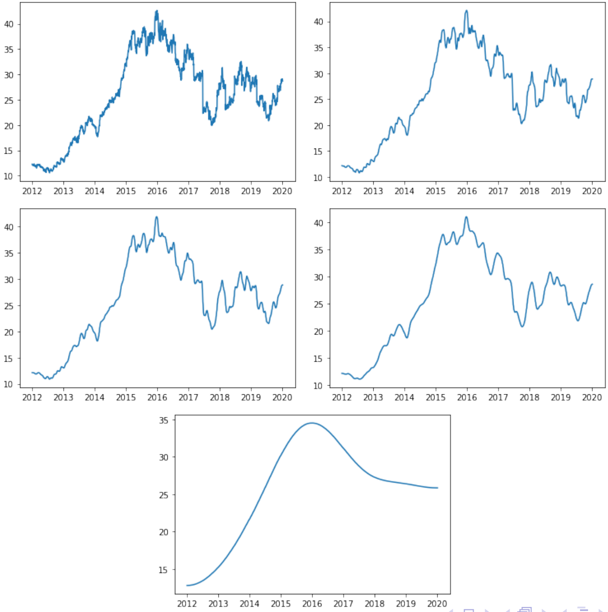


Figure: Left: Gaussian with mean 0, assigns higher weights to neighboring points. Right: Application in image processing

- Caveat: Smoothing of a time series should be causal, because we should never use future values of y. May not make sense

Apply smoothing on KR.



- Compute first derivative as:
$$y'[i] = y[i+1] - y[i]$$
- Compute second derivative as:
$$y''[i] = y'[i+1] - y'[i]$$
- Parameters to tune: filter window b

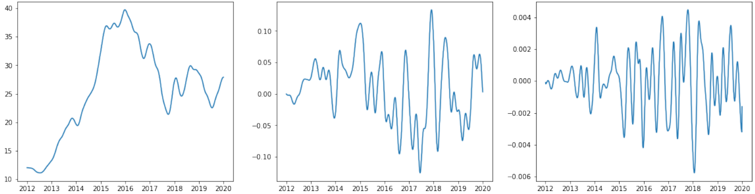
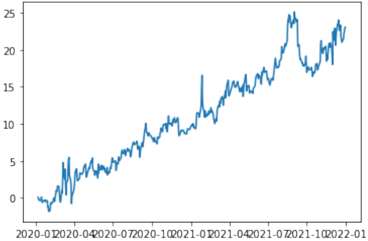


Figure: $b = 5$, original, first derivative, second derivative

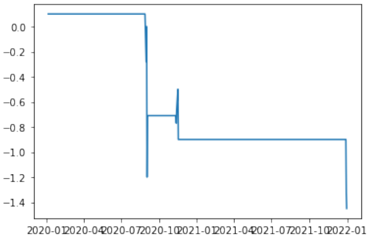
Trend following

- Buy whenever the estimated first derivative is > 0, sell whenever < 0
- Final chosen $b = 50$
- Final test Sharpe: 0.2276



Contrarian

- Buy whenever we detect a local min; sell whenever we detect a local max
- Parameters to tune: b smoothing window, ϵ tolerance
- Final chosen $b = 50$, $\epsilon = 0.001$



Contrarian did not perform very well

Polynomial Regression

- Goal: smoothing allows us to approximate derivatives more conveniently, but still an approximation $\rightarrow$ Can have analytic formulae?
- Transform price data using monomials of different degrees, arrange them in a matrix and fit in squared error
- Want to fit: $y = f(x)$ with data $(x_i, y_i)_{i=1}^{N}$. Model:

$$f(x) = a_0 + a_1 x + \cdots + a_k x^k = \begin{bmatrix} 1 & x & \cdots & x^k \end{bmatrix} \cdot \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_k \end{bmatrix}$$

- Want to match at least at the data points $\{y_i\}_{i=1}^{N}$, we minimize:

$$\sum_{i=1}^{N} (y_i - f(x_i))^2$$

- Arrange into matrix, we have:

$$f(X) = \underbrace{\begin{bmatrix} 1 & x_1 & \cdots & x_1^k \\ 1 & x_2 & \cdots & x_2^k \\ \vdots & \cdots & \ddots & \vdots \\ 1 & x_N & \cdots & x_N^k \end{bmatrix}}_{=:\Phi(X),\text{"feature matrix"}} \underbrace{\begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_k \end{bmatrix}}_{=:\alpha}$$

and solve for coefficients α. The optimal solution can be analytically calculated by setting derivative to 0: $\hat{\alpha} = (\Phi^T \Phi)^{-1} \Phi^T y$.

- After obtaining α, we have a polynomial that fits our data which we can exactly differentiate and predict trends.
- This is available in Python as `numpy.polyfit` or `sklearn.preprocessing.Polynomialfeatures`

- We test trend following view using smoothed version. This time we can take analytic derivatives.
- Parameters tuned: highest degree of polynomial $k \in [4, 10]$.
- Final degree found: 9; Sharpe ratio: 0.3553

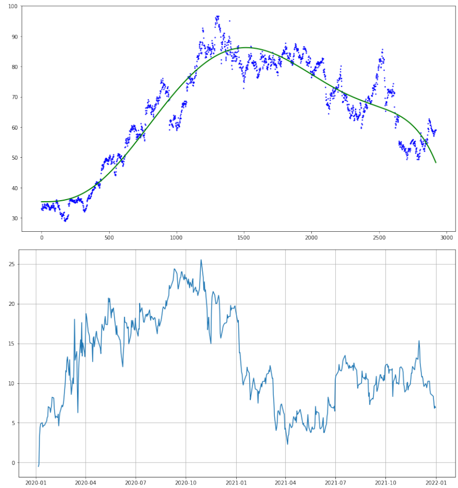


Figure: WBA trading cumulative PNL with degree 9 polynomial