Understanding Asset Return Process

Jacques Nel

York University

April 30, 2020

Our problem

Develop a learning-to-rank investment strategy following the approach of (Liu et al., 2017)

1. Given our universe of n stocks, pick top k
2. Train model on 2 years
3. Test performance on 3rd year
4. Report performance metrics: top-k rank accuracy, average return, volatility, sharpe ratio, and max drawdown
5. Compare with `S&P500` index

Achieving stationarity

Neural network performance is dependent on stationarity of training data
Intuitively:

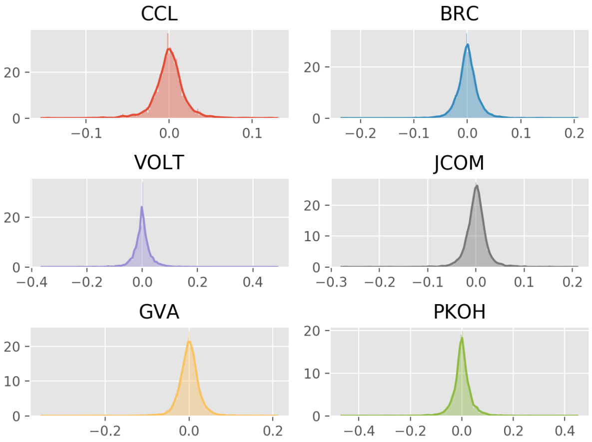
1. Each example should look similar to the next, minus the dynamics we are trying to model
2. Numerical performance implications

In time series problems (including regression with ARIMA or GARCH), apply differencing scheme to make data stationary.
Log-returns are the best choice in financial time series, ie., given a time series $(s_t) \in \mathbb{R}$, for $0 < t \in \mathbb{Z}$:

$$x_t = \log s_t - \log s_{t-1}$$

Exploring our data

Figure: Log-returns for 6 stocks



Some observations:

1. Log-returns maybe folow Laplace distribution
2. Distributions have heavy tails

Dilated Temporal Convolution Network (TCN) ie. WaveNet

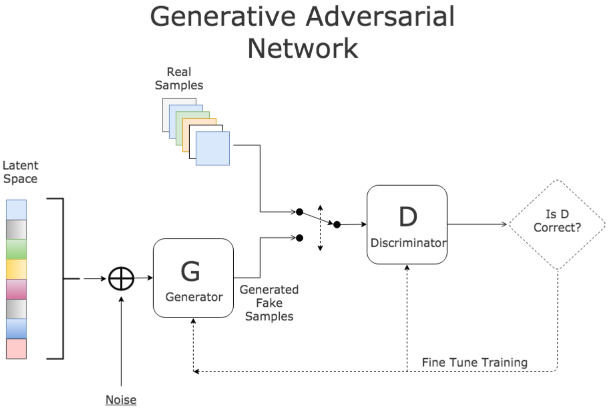
(van den Oord et al., 2016) introduces high performance generative model for audio waveforms

1. Overcomes training problems of LSTMs, GRUs, and other RNN architectures
2. No vanishing gradient
3. Massive improvement in performance and memory usage
4. Attains very long look-back (receptive field) using dilated convolutions

Dilated Temporal Convolution Network (TCN) ie. WaveNet

Generative Adverserial Network (GAN)

GANs are universal generators (analgolous to universal approximation theorem)



Lambert W transformation

Lambert's W function is the inverse of $z = u\exp(u)$, ie. the function that satisfies

$$W(z)\exp(W(z)) = z \tag{1}$$

Let U be a continous RV with cdf $F_U(u|\boldsymbol{\beta})$, pdf $f_U(u|\boldsymbol{\beta})$, and parameter vector $\theta = (\boldsymbol{\beta}, \delta)$, then

$$Z = U\exp\left(\frac{\delta}{2}U^2\right), \quad \delta \in \mathbb{R} \tag{2}$$

is a non-central, non-scaled heavy tail Lambert $W \times F_X$ RV with parameter vector $\theta = (\boldsymbol{\beta}, \delta)$, where δ is the tail parameter.

Lambert W transformation

For a continuous location-scale family RV X $F_X(x|\boldsymbol{\beta})$ define a location-scale heavy-tailed Lambert $W \times F_X$ RV

$$Y = \left\{ U \exp\left(\frac{\delta}{2} U^2\right) \right\} \sigma_x + \mu_x, \quad \delta \in \mathbb{R} \tag{3}$$

Let X $F_X(x|\boldsymbol{\beta})$ be continuous scale-family RV, with standard deviation σ_x, let $U = X/\sigma_x$, then

$$Y = X \exp\left(\frac{\delta}{2} U^2\right), \quad \delta \in \mathbb{R}, \tag{4}$$

is a heavy-tailed Lambert $W \times F_X$ RV with parameter $\theta = (\boldsymbol{\beta}, \delta)$. Let $\tau := (\mu_x(\boldsymbol{\beta}), \sigma_x(\boldsymbol{\beta}))$ define transformation eq:trans.

"Gaussianize" heavy-tailed data

The inverse transformation of eq:trans is

$$W_\tau(Y) := W_\delta\left(\frac{Y-\mu_x}{\sigma_x}\right)\sigma_x + \mu_x = U\sigma_x + \mu_x = X \tag{5}$$

with

$$W_\delta(z,\delta) := (z)\left(\frac{W(\delta z^2)}{\delta}\right)^{1/2}. \tag{6}$$

$W_\delta(z)$ is bijective for $\forall \delta > 0$ and $\forall z \in \mathbb{R}$.

MSE

Angqi Liu, Qiang Song, and Steve Y. Yang. Stock portfolio selection using learning-to-rank algorithms with news sentiment. *s*, 2017.

Aaron van den Oord, Sander Dieleman, Heiga Zen, Karen Simonyan, Oriol Vinyals, Alex Graves, Nal Kalchbrenner, Andrew Senior, and Koray Kavukcuoglu. Wavenet: A generative model for raw audio. *s*, 2016.