Deep Learning in Asset Pricing

Wu Zhu (Presenter)
University of Pennsylvania

December 5, 2019

Motivation

Deep learning is nothing but a powerful non-parametric (functional) approximation.

Incorporate these powerful functional approximation tools into finance and economics, but still remain basic economic intuitions

Two recent papers - will emphasize how to link economic sense with our toolbox

Background in Economics

Suppose we have N stocks or assets,

[1] Excess return vector: $R_t^e = (R_{1t}^e, ..., R_{Nt}^e)'$ with observations $t = 1, 2, ... T$.

[2] Economic constraint - agents try to make arbitrage using their information. As a result, there would be no arbitrage in equilibrium, i.e., there exists

$$E_t M_{t,t+1} R_{i,t+1}^e = 0 \Longrightarrow E_t R_{i,t+1}^e = -\frac{Cov_t(M_{t+1}, R_{t+1}^e)}{Var_t(M_{t+1})} \frac{Var_t(M_{t,t+1})}{E_t M_{t,t+1}}$$

with $E_t[.]$ the conditional expectation operator conditional on information up till period t.

$M_{t,t+1}$ is the SDF (Stochastic Discount Factor) measuring the relative prices of one dollar in different states

Background in Economics

Suppose agents optimize their portfolios, then there exist a time-varying vector w_t such that

$$M_{t,t+1} = 1 - w_t^T R_{t+1}^e$$

where w_t is measurable with respect to information set uptill t.

$$E[[1 - w_t^T R_{t+1}^e] R_{t+1}^e g(I_t, I_{i,t})] = 0$$

for $\forall$ measurable function g with respect to information set up till t

I_t are all available macro observable variables, and I_{it} are firms' individual characteristics uptill period t.

Deep Learning in Asset Pricing (Chen,Pelger, and Zhu, 2019)

Deep Learning in Asset Pricing (Chen, Pelger, and Zhu,2019)

Deep Learning in Asset Pricing (Chen,Pelger, and Zhu, 2019)

How do they estimate time varying w_t? Generative Adversarial Networks

$$\min_{w}\max_{g}\frac{1}{N}\sum_{j=1}^{N}\|\,\mathbb{E}\Big[\big(1-\sum_{i}w(I_t,I_{t,i})R^e_{t+1,i}\big)R^e_{t+1,j}\,g(I_t,I_{j,t})\Big]\,\|^2$$

Empirically, define loss function

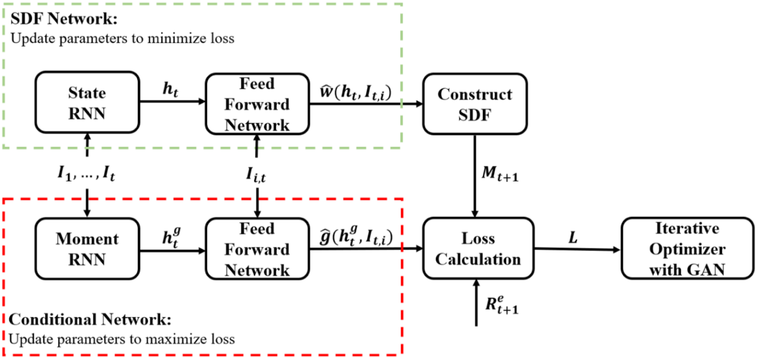
$$L(w,g,I_t,I_{t,i})=\frac{1}{N}\sum_{i=1}^{N}\frac{T_i}{T}\,\|\,\frac{1}{T_i}\sum_{t_i}M_{t+1}R^e_{t+1,i}\,g(I_t,I_{t,i})\,\|^2$$

Deep Learning in Asset Pricing (Chen,Pelger, and Zhu, 2019)

1. $w(I_t, I_{i,t})$ - approximate using RNN
2. $g(I_t, I_{i,t})$ - approximate using RNN
3. $w(I_t, I_{i,t})$ interact with $g(I_t, I_{i,t})$ in adversarial pattern - Given $w(I_t, I_{i,t})$, $g(I_t, I_{i,t})$ maximizes the loss, given $g(I_t, I_{i,t})$, $w(I_t, I_{i,t})$ minimizes the loss

Deep Learning in Asset Pricing (Chen,Pelger, and Zhu, 2019)

Figure 1. Model Architecture



Model architecture of GAN (Generative Adversarial Network) with RNN (Recurrent Neural Network) with LSTM cells.

Deep Learning in Asset Pricing (Chen,Pelger, and Zhu, 2019)

Table III Performance of Different SDF Models

	SR			EV			Cross-Sectional R^2		
Model	Train	Valid	Test	Train	Valid	Test	Train	Valid	Test
LS	1.80	0.58	0.42	0.09	0.03	0.03	0.15	0.00	0.14
EN	1.37	1.15	0.50	0.12	0.05	0.04	0.17	0.02	0.19
FFN	0.45	0.42	0.44	0.11	0.04	0.04	0.14	-0.00	0.15
GAN	2.68	1.43	0.75	0.20	0.09	0.08	0.12	0.01	0.23

Monthly Sharpe Ratio (SR) of the SDF factor, explained time series variation (EV) and cross-sectional mean R^2 for the GAN, FFN, EN and LS model.

Deep Learning in Asset Pricing (Chen,Pelger, and Zhu, 2019)

Table IV SDF Factor Risk Measures

	SR			Max Loss			Max Drawdown		
Model	Train	Valid	Test	Train	Valid	Test	Train	Valid	Test
FF-3	0.27	-0.09	0.19	-2.45	-2.85	-4.31	7	10	10
FF-5	0.48	0.40	0.22	-2.62	-2.33	-4.90	4	3	7
LS	1.80	0.58	0.42	-1.96	-1.87	-4.99	1	3	4
EN	1.37	1.15	0.50	-2.22	-1.81	-6.18	1	3	5
FFN	0.45	0.42	0.44	-3.30	-4.61	-3.37	6	3	5
GAN	2.68	1.43	0.75	0.38	-0.28	-5.76	0	1	5

Sharpe Ratio, maximum 1-month loss and maximum drawdown of the SDF factor portfolios. We include the mean-variance efficient portfolio based on the 5 Fama-French factors.

Autoencoder Asset Pricing Models (Gu, Kelly, and Xiu, 2019)

Autoencoder Asset Pricing Models - 2019

Autoencoder Asset Pricing Models (Gu, Kelly, and Xiu, 2019)

Economic Backgrounds

Consider a standard factor model. Suppose the SDF takes factor structure, we have

$$r_{i,t+1} = \beta'_{i,t} f_{t+1} + \epsilon_{t+1}$$

where $\beta_{i,t}$ is the loading of stock i on factor f_{t+1}. Usually, both the loadings and factors are unobservable.

with economic constraint,

$$E_t[\epsilon_{i,t+1}] = 0, Cov_t(f_{t+1}, \epsilon_{i,t+1}) = 0$$

Autoencoder Asset Pricing Models (Gu, Kelly, and Xiu, 2019)

Loss function

$$L = \frac{1}{NT} \sum_{t=1}^{T} \sum_{i=1}^{N} (r_{it+1} - \beta'_{i,t} f_{t+1})^2$$

How to approximate for $\beta_{it} = \beta(z_{it})$ - use deep neural net

How to approximate for factors f_{t+1} - use autoencoder, that is, initially use r_{t+1}, and take a dimension reduction to recover lower dimensional factors.

Autoencoder Asset Pricing Models (Xiu et al 2019)

$$z_{i,t-1}^{(0)} = z_{i,t-1},$$

$$z_{i,t-1}^{(l)} = g\left(b^{(l-1)} + W^{(l-1)} z_{i,t-1}^{(l-1)}\right), \quad l = 1, \ldots, L_\beta,$$

$$\beta_{i,t-1} = b^{(L_\beta)} + W^{(L_\beta)} z_{i,t-1}^{(L_\beta)}.$$

Autoencoder Asset Pricing Models (Xiu et al 2019)

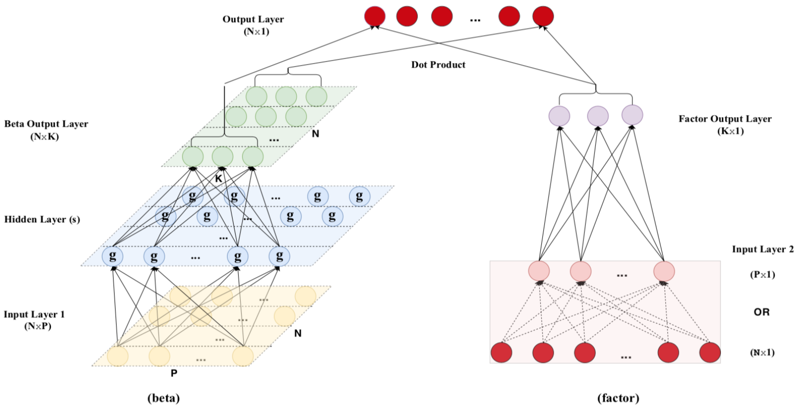
$$r_t^{(0)} = r_t,$$

$$r_t^{(l)} = \widetilde{g}\left(\widetilde{b}^{(l-1)} + \widetilde{W}^{(l-1)} r_t^{(l-1)}\right), \quad l = 1, \ldots, L_f,$$

$$f_t = \widetilde{b}^{(L_f)} + \widetilde{W}^{(L_f)} r_t^{(L_f)}.$$

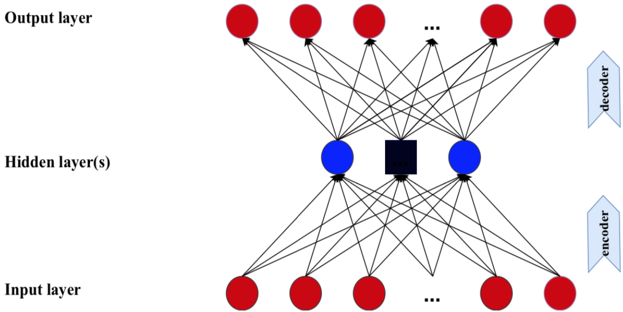
Autoencoder Asset Pricing Models (Xiu et al 2019)

Figure 2: Conditional Autoencoder Model



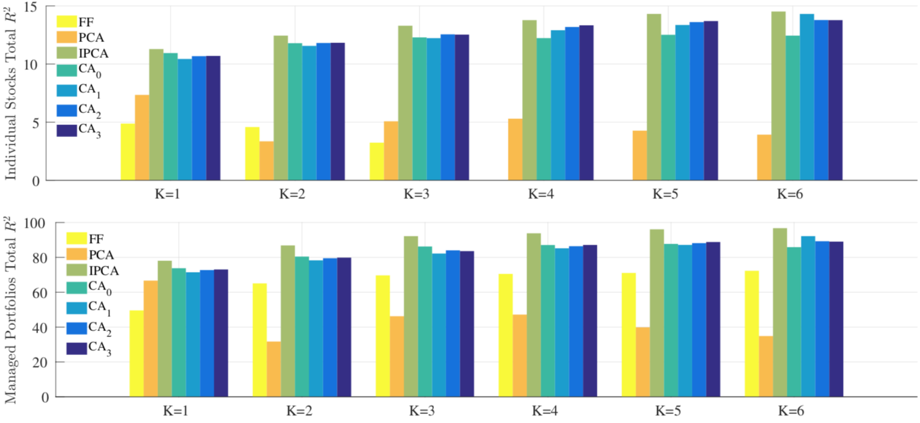
Autoencoder Asset Pricing Models (Xiu et al 2019)

Figure 1: Standard Autoencoder Model



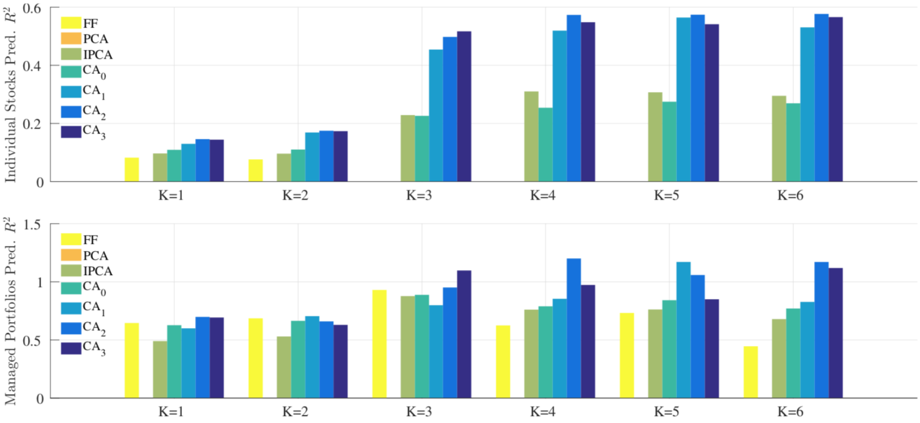
Note: This figure describes a standard autoencoder with one hidden layer. The output and input layers are identical, while the hidden layer is a low dimensional compression of inputs variables into latent factors, which can be expressed as weighted linear combinations of input variables.

Autoencoder Asset Pricing Models (Xiu et al 2019)



Note: In this table, we report the out-of-sample total $R^2(\%)$ for individual stocks r_t and managed portfolios x_t using observable factor models (FF), PCA, IPCA, and conditional autoencoders CA_0 through CA_3. In all cases, the number of factors K varies from 1 to 6.

Autoencoder Asset Pricing Models (Xiu et al 2019)



Note: In this table, we report the out-of-sample predictive $R^2(\%)$ for individual stocks r_t and managed portfolios x_t using observable factor models (FF), PCA, IPCA, and conditional autoencoders CA_0 through CA_3. In all cases, the number of factors K varies from 1 to 6.

Autoencoder Asset Pricing Models (Xiu et al 2019)

Table 3: Out-of-Sample Sharpe Ratios of Long-Short Portfolios

	K					
Equal-Weight	1	2	3	4	5	6
FF	-0.66	-0.85	-0.40	-0.30	0.36	-0.21
PCA	0.28	0.09	0.13	-0.08	-0.12	0.15
IPCA	0.20	0.19	1.26	2.16	2.31	2.25
CA_0	0.23	0.32	1.34	1.87	2.10	2.18
CA_1	0.30	0.39	2.12	2.63	2.67	2.60
CA_2	0.30	0.38	2.16	2.64	2.68	2.63
CA_3	0.31	0.38	2.19	2.57	2.57	2.59
	K					
Value-Weight	1	2	3	4	5	6
FF	-0.82	-1.13	-0.69	-0.60	0.18	-0.53
PCA	0.12	-0.18	0.05	-0.10	-0.30	-0.08
IPCA	-0.15	-0.07	0.59	0.81	1.05	0.96
CA_0	-0.11	-0.03	0.41	0.81	0.83	0.88
CA_1	-0.03	0.11	0.91	1.30	1.48	1.40
CA_2	-0.03	0.08	0.92	1.39	1.45	1.53
CA_3	-0.02	0.08	1.09	1.41	1.34	1.51

Note: In this table, we report annualized out-of-sample Sharpe ratios for long-short portfolios using Fama-French models (FF), a vanilla factor model (5), and a variety of autoencoders, A_0, A_1, A_2, A_3, based on (9), respectively, where the number of factors in (5) or the number of neurons in the hidden layer on the right-hand side of (9), K, varies from 1 to 6.