



# Machine Learning for Prediction

Iris He (Lingyun He): he@rams.colostate.edu and Rui Tang: tangrui@rams.colostate.edu

Mentor: Edwin Chong: Edwin.Chong@colostate.edu

Department of Electrical and Computer Engineering - Colorado State University

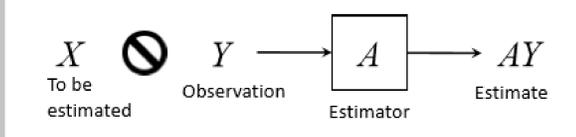


## Abstract

- ▶ For this project, we use machine learning to predict economic data. Machine learning is a way to identify data patterns and use them to automatically make predictions or decisions.
- ▶ A large part of the project is to explore different mathematical methods to predict future economic data from the past economic data. We implemented different methods by computer program using MATLAB for data fitting. We also compared the accuracy of different models. And then application is also an important part. We need to apply their results to reality and discuss efficiency and feasibility. Our goal is to find the most accurate, feasible and effective machine learning model for predicting future economic data.
- ▶ In the project, we use mean squared error to compare the performance of different prediction methods. In statistics, the mean squared error (MSE) of an estimator (a procedure for estimating an unobserved quantity) measures the average of the squares of the errors—that is, the average squared difference between the estimated values and the actual values. We focus on the family of linear minimum mean-squared error estimation problems with constraints on the estimators.
- ▶ The research is under direction of Dr.Edwin Chong.

## Basic Math Background

- ▶ Suppose that  $X$  is an  $N$ -variate complex random vector and  $Y$  is an  $M$ -variate complex random vector. Assume that both have zero mean. Thus, for each of  $X$  and  $Y$ , the covariance and correlation are equal. We wish to estimate  $X$  from  $Y$  using a linear estimator  $A$  ( $N \times M$  matrix) such that the estimate  $AY$  minimizes the mean-squared error  $E[\|AY - X\|^2]$ , where  $E$  represents either expectation (estimated using the empirical mean from data), and  $\|\cdot\|$  is the standard complex 2-norm.



- ▶ In other words, the purpose of our work is to solve the optimization problem:
 
$$\text{minimize } E[\|AY - X\|^2]$$
- ▶ To make the covariance of the data stationary, we assign exponential weights  $(\gamma^0, \gamma^1, \dots, \gamma^k)$  to emphasize the most recent periods of data.

## Matrix

Form data matrix  
We have 2000 raw data points and set the first data point as  $P(1)$ . The last data point as  $P(2000)$ . Choose 2000 data point to do this, and choose  $M = 450$ ,  $N = 50$ .

$$Z = \begin{pmatrix} P(1) & P(2) & \dots & P(500) \\ P(2) & P(3) & \dots & P(501) \\ \vdots & \vdots & \ddots & \vdots \\ P(1500) & P(1501) & \dots & P(1999) \end{pmatrix}$$

1. Scale (Normalization)  
Choose the last component of  $Y$ , then all values in the data matrix  $Z$  will be divided by the last component of  $Y$  to do normalization. (The components of  $Y$  are ordered chronologically. The last component of  $Y$  is the price value for the most recent day.)

2. Center (Subtract mean)

Let

$$\bar{x} = \frac{1}{k} \sum_{i=1}^k x_i$$

be the average vector. We subtract  $\bar{x}$  from each  $x_i$ . We further assume that the vectors  $x_1, x_2, \dots, x_k$  are drawn from the same underlying distribution.

3. Covariance

$\gamma = 0.98$  was selected. The sample covariance matrix is calculated as:

$$\Sigma_{z,z} = \frac{1-\gamma}{1-\gamma^{k+1}} Z' \text{diag}(\gamma^0, \gamma^1, \dots, \gamma^k) Z$$

$$\Sigma_{z,z} = \begin{pmatrix} \Sigma_{y,y} & \Sigma_{y,x} \\ \Sigma_{x,y} & \Sigma_{x,x} \end{pmatrix}$$

$$\Sigma_{y,y} = C_Y, \Sigma_{x,x} = C_X$$

$$\Sigma_{x,y} = C_{XY}, \Sigma_{y,x} = C_{YX}$$

## Prediction Methods

1. Wiener filter

$$A_{Wiener} = C_{XY} C_Y^{-1}$$

2. Low-rank Wiener filter

$$C_{XY} C_Y^{-\frac{1}{2}} = USV'$$

$$A_{LRW} = C_{XY} C_Y^{-\frac{1}{2}} V_L V_L' C_Y^{-\frac{1}{2}}$$

3. PC Wiener Filter

let  $L < M + N$  be such that the first  $L$  singular values account for the bulk of the sum of the squared singular values (say 0.90).

$$C_Y^{PC} = V_L S_L V_L'$$

$$A_{PC} := C_{XY} C_Y^{-PC} = C_{XY} (C_Y^{PC})^{-1}$$

4. Cross-Spectral Wiener Filter

Develop cross-spectral Wiener filter and call it as  $A_{CS}$ . Suppose that we order the columns of  $V$  and the elements of  $S$  according to the cross-spectral powers, from largest to smallest. Then we can get the different covariance. Then we can use the same formula of  $A_{PC}$  to compute  $A_{CS}$

The quantity  $\|\phi_i\|$  is called the  $i$ th cross-spectral power of  $X$  and  $Y$ .  $\phi_i$  is defined in as the  $i$ th column of  $C_{XW}$ , and  $\|\phi_i\|$  is its Euclidean norm. Order the cross-spectral powers, from largest to smallest.

5. Joint-PC Wiener Filter

$$Z = \begin{pmatrix} X \\ Y \end{pmatrix}, V_Z = \begin{pmatrix} V_X \\ V_Y \end{pmatrix}, V_Y = (V_{Y,L} \ V_{Y,\bar{L}})$$

$$A_{JPC} = C_{XY} V_{Y,L} (V_{Y,L}' C_Y V_{Y,L})^{-1} V_{Y,L}'$$

6. Least-Squares Joint-PC Wiener Filter

$$V_X = (V_{X,L} \ V_{X,\bar{L}})$$

$$A_{LSJPC} = V_{X,L} (V_{Y,L}' V_{Y,L})^{-1} V_{Y,L}'$$

## Mean Squared Error (MSE) and Conclusion

Normalized the mean squared error:

$$RMSE = MSE / norm = \sqrt{E[\|AY - X\|^2]} / \sqrt{E[\|X\|^2]}$$

1. Dow Jones Industrial Average  
To test the methods, we first used data from the Dow Jones Industrial Average. The data is from 07-May-2013 to 15-April-2021.

$$RMSE_{Wiener} = 1.9414e + 03$$

$$RMSE_{LRW} = 1.9414e + 03$$

$$RMSE_{PC} = 1.0023, RMSE_{CS} = 1.0002$$

$$RMSE_{JPC} = 0.9970, RMSE_{LSJPC} = 1$$

2. Apple Inc.

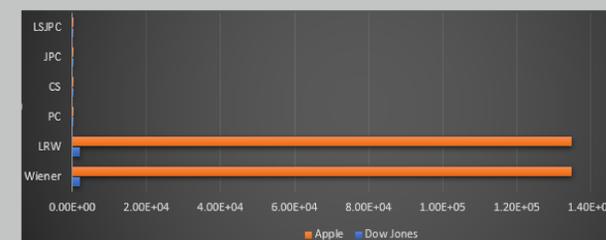
The second data set was for Apple. The data is from 07-May-2013 to 15-April-2021.

$$RMSE_{Wiener} = 1.3492e + 05$$

$$RMSE_{LRW} = 1.3492e + 05$$

$$RMSE_{PC} = 1.0066, RMSE_{CS} = 1.0000$$

$$RMSE_{JPC} = 0.9984, RMSE_{LSJPC} = 1$$



## Conclusion

From the results above, we can see that the results of Wiener filter and the low-rank Wiener filter are larger, compared with the results of PC Wiener Filter, Cross-Spectral Wiener Filter, Joint-PC Wiener Filter and Least-Squares Joint-PC Wiener Filter. The last four models have better performance than the first two filters.