Data Analysis and Optimization of ALM Problem

Presented by : Gopal Kalpande

Roll No.:17MCMI10

Project Guide : Dr. V. N. Sastry

Date: 27th August 2019

Outline

- Objective of Project
- Risk Associated with Banking
- Introduction to ALM
- Components of Bank Balance Sheet
- Proposed Solutions
  - Simulation and Analysis of ALM
  - Linear optimization of Liquidity Risk Management
  - LSTM implementation for Prediction of Stock Price
- Conclusion
- Future Work
- References

Chapter 1

Introduction

Objective of Project

Creation of Strategies to Stabilize Financial Networks and to Improve Profitability of them from various risks.

The aim was to study the LPP optimization method and Deep Learning for finding the solution to risks associated with banking.

Risks Associated with Banking

- Credit Risk
- Operational Risk
- Liquidity Risk
- Market Risk
    - Equity risk
    - Interest Rate Risk
    - Foreign Exchange Risk
    - Commodity risk

Introduction to ALM

What is ALM

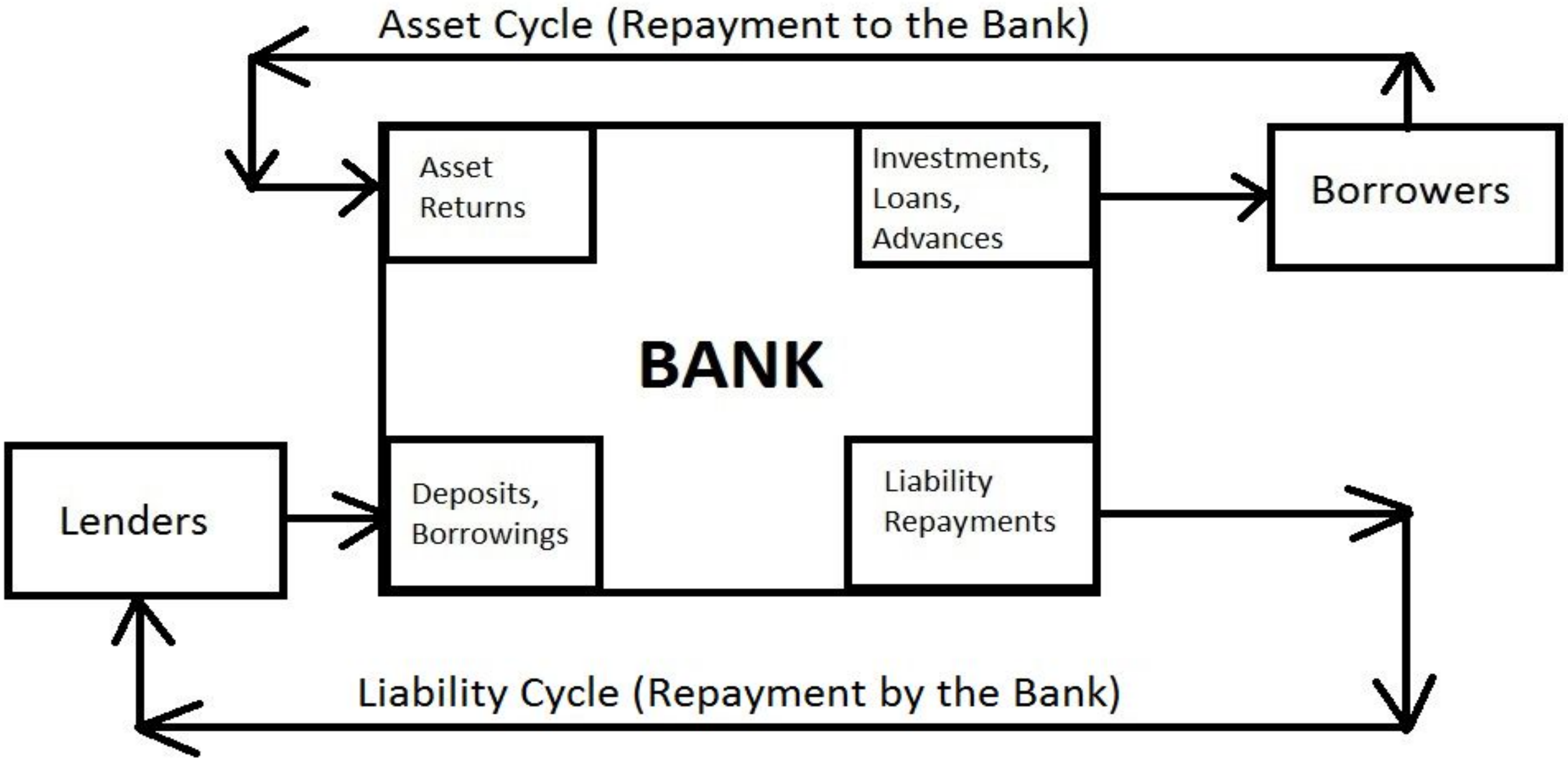
Periodic monitoring of risk exposures involving collecting and **analysing** information

Ability to anticipate, **forecast** and act so as to structure bank’s business to profit

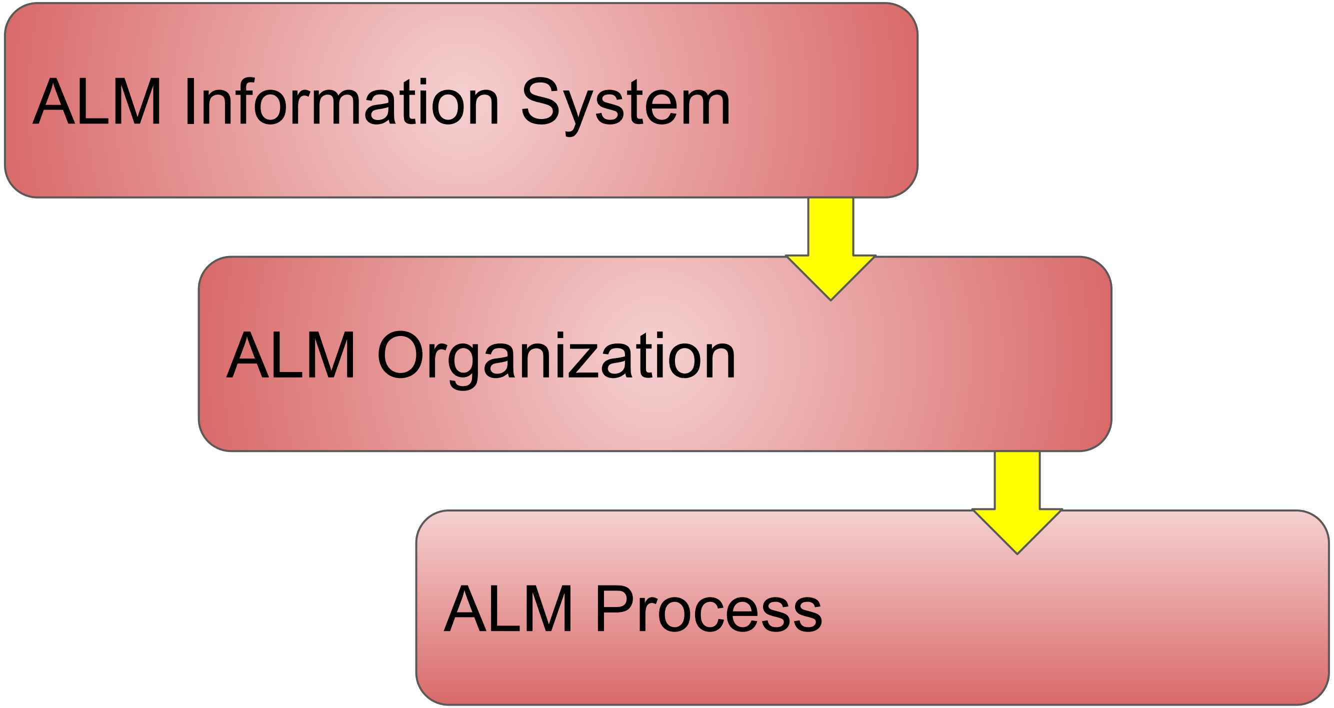
Altering A & L portfolio in a dynamic way to **manage risks**

Involves judgement and **decision making**

ALM Cycle of a Bank



Three Pillars of ALM



Maturity Buckets

1) Next day
2) 2 to 14 days
3) 15 to 28 days
4) 29 days and upto 3 months
5) Over 3 months and upto 6 months
6) Over 6 months and upto 1 year
7) Over 1 year and upto 3 years
8) Over 3 years and upto 5 years
9) Over 5 years and upto 7 years
10) Over 7 years and upto 10 years
11) Over 10 years.

Components of a Bank Balance sheet

Liabilities	**Assets**
● Capital ● Reserve Funds & Surplus ● Deposits ● Borrowings ● Foreign Currency Liability	● Cash & Balances with RBI ● Bal. With Banks & Money at Call and Short Notices ● Investments ● Fixed Assets, Loans and Advances ● Foreign Currency Assets

Chapter 2

Simulation and Analysis of ALM

2.1 Simulation and Analysis of ALM

Asset :

1. Term Loans

Liabilities :

1. Deposits

    (i) Savings Bank Deposits

2. Borrowings

    (i) Call and Short Notice

    (ii) Inter-Bank (Term)

2.2 Dataset Description

Name : 1999 Czech Financial Dataset - Real Anonymized Transactions

Data from a real Czech Bank from 1999.

-relation transaction (1056320 objects in the file TRANS.ASC) - each record describes one transaction on an account,

-relation loan (682 objects in the file LOAN.ASC) - each record describes a loan granted for a given account,

2.3 Description of Transaction File

```
## # A tibble: 6 x 10
##   trans_id account_id   date type  operation amount balance k_symbol bank
##      <dbl>      <dbl>  <dbl> <chr> <chr>      <dbl>   <dbl> <chr>    <chr>
## 1   695247       2378 930101 PRIJ~ VKLAD        700     700 <NA>     <NA>
## 2   171812        576 930101 PRIJ~ VKLAD        900     900 <NA>     <NA>
## 3   207264        704 930101 PRIJ~ VKLAD       1000    1000 <NA>     <NA>
## 4  1117247       3818 930101 PRIJ~ VKLAD        600     600 <NA>     <NA>
## 5   579373       1972 930102 PRIJ~ VKLAD        400     400 <NA>     <NA>
## 6   771035       2632 930102 PRIJ~ VKLAD       1100    1100 <NA>     <NA>
## # ... with 1 more variable: account <dbl>
```

2.4 Description of Loan File

```
## # A tibble: 6 x 7
##   loan_id account_id   date amount duration payments status
##     <dbl>      <dbl>  <dbl>  <dbl>    <dbl>    <dbl> <chr>
## 1    5314       1787 930705  96396       12     8033 B
## 2    5316       1801 930711 165960       36     4610 A
## 3    6863       9188 930728 127080       60     2118 A
## 4    5325       1843 930803 105804       36     2939 A
## 5    7240      11013 930906 274740       60     4579 A
## 6    6687       8261 930913  87840       24     3660 A
```

2.5 Different Scenarios for Simulation

- Different Maturity Buckets as referred on slide 10
- Using only Loan file
- Using only Transaction file
- By combining the data of Loan and Transaction files

2.6

Algorithm 2 Pseudo Code for Computation of Maturity Table for ALM

1: Initialization:
Start ← [The starting day of each time bucket]
End ← [The ending day of each time bucket]
i ← 0
max ← length(Start)
2: *while(i < max)*:
3:
$$Liability = \Sigma_{j=Start[i]}^{End[i]} deposit\_amount[j] \tag{2.1}$$
4:
$$Asset[i] = \Sigma_{j=Start[i]}^{End[i]} loan\_amount[j] \tag{2.2}$$
5:
$$Withdrawal[i] = \Sigma_{j=Start[i]}^{End[i]} withdrawal\_amount[j] \tag{2.3}$$
6:
$$Mismatch[i] = Asset[i] - Withdrawal[i] \tag{2.4}$$
7:
$$\%\_Mismatch[i] = \frac{Mismatch[i]}{Liability[i]} \tag{2.5}$$
8: Cumulative % Mismatch = Cumulative_Sum(%_Mismatch)

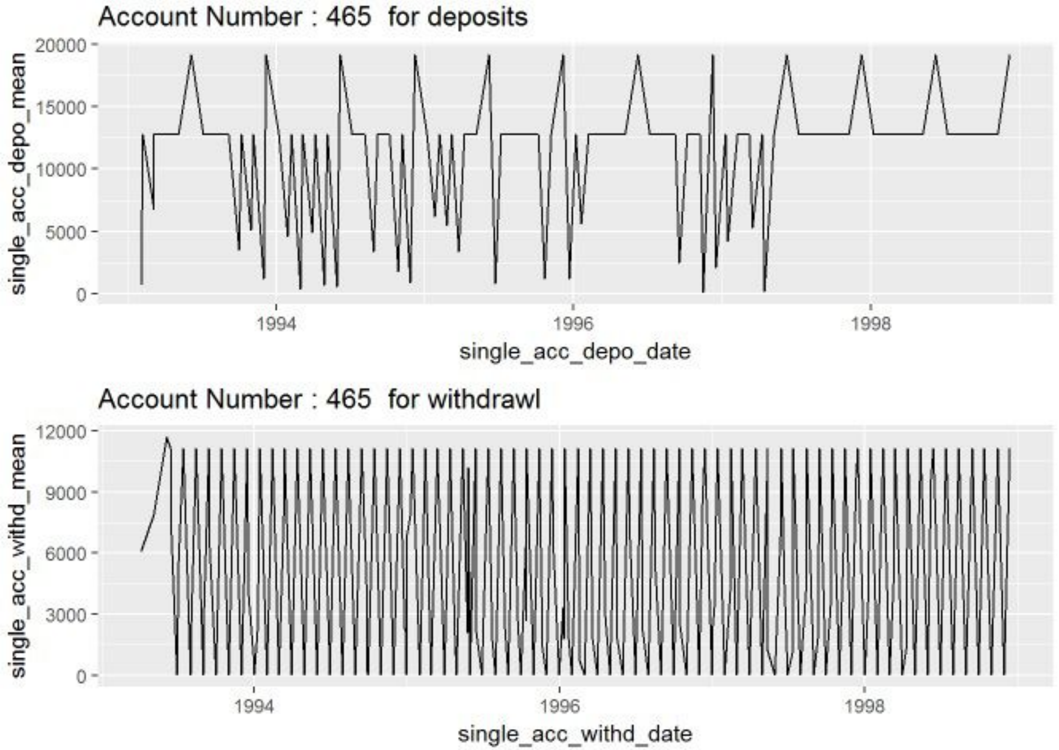
2.7 Results of the ALM Simulation

B.N.	#Start	#End	Mismatch	% Mismatch	Cumulative % Mismatch
Bucket#1	1	1	3200	-100.00000	-100.0000
Bucket#2	2	7	28964	-100.00000	-200.0000
Bucket#3	8	14	268887	-100.00000	-300.0000
Bucket#4	15	28	296622	-100.00000	-400.0000
Bucket#5	29	90	6348345	-99.99665	-499.9966
Bucket#6	91	180	23510757	-99.98661	-599.9833
Bucket#7	181	265	89174600	-85.65359	-685.6368
Bucket#8	366	1095	669010939	-51.31571	-736.9526
Bucket#9	1096	1825	1458973798	-55.54921	-792.5018
Bucket#10	1826	2179	942042305	-50.90293	-843.4047

2.8 Results of the ALM Simulation

B.N.	#Start	#End	Liability	Withdrawal	Assets
Bucket#1	1	1	3200	0	0
Bucket#2	2	7	28964	0	0
Bucket#3	8	14	268887	0	0
Bucket#4	15	28	296622	0	0
Bucket#5	29	90	6348345	1166308	0
Bucket#6	91	180	23510757	13085553	0
Bucket#7	181	265	89174600	64642130	2619276
Bucket#8	366	1095	669010939	577121716	26724276
Bucket#9	1096	1825	1458973798	1261537518	48700920
Bucket#10	1826	2179	942042305	832618418	25217268

2.9 Post Result Analysis of Data



2.10 The formula for finding the assets from each saving accounts can be given as: {Contribution}

$$ASA = \{AAD - AAW\}_{overgiventimebucket} \tag{2.6}$$

Where,

- ASA: Assets from Saving Account
- AAD: Average Amount Deposited
- AAW: Average Amount Withdrawal

2.11 Average balance of randomly selected saving account

```
##    account_nos acc_depo_mean acc_withd_mean differences
## 1         2378     14433.676      10355.634    4078.042
## 2          465      6371.113       5145.206    1225.906
## 3         5270     23756.868      10567.704   13189.164
## 4         1019     10565.604       3661.526    6904.078
## 5         1637      2906.244       1367.747    1538.497
## 6          485     14473.943       6214.987    8258.956
## 7         3367     12334.281       5660.907    6673.374
## 8         2486     17729.990       9340.120    8389.870
## 9         3678      5079.980       3149.784    1930.196
## 10        1127     12698.727       6580.909    6117.818
```

Chapter 3

Linear optimization of Liquidity Risk Management

3.1 Linear Optimization of Liquidity Risk Management

Objective Function:

$$Maximize G_b = \sum_{a=1}^{k-1} \sum_{c=1}^{m} \{Y_c^{a,b} + Y_c^{a,b} * IA_c^{a,b} * TA_c^{a,b}\} - \sum_{a=1}^{k-1} \sum_{d=1}^{n} \{X_d^{a,b} + X_d^{a,b} * IL_d^{a,b} * TL_d^{a,b}\}$$

c - Asset/Liability number

a - investment time bucket

b - maturity time bucket

G - Gap value (Assumed to be given)

3.2

Algorithm 3 The Algorithm for LPP Optimization

1: Objective Function:

$$MaximizeProfit = \Sigma_{j=1}^{m} Asset_j - \Sigma_{i=1}^{n} Liability_i \tag{3.1}$$

which can further be simplified as

$$MaximizeG_b = \sum_{a=1}^{k-1}\sum_{c=1}^{m}\{Y_c^{a,b}+Y_c^{a,b}*IA_c^{a,b}*TA_c^{a,b}\}-\sum_{a=1}^{k-1}\sum_{d=1}^{n}\{X_d^{a,b}+X_d^{a,b}*IL_d^{a,b}*TL_d^{a,b}\} \tag{3.2}$$

2: Assumption:
a. No amount is expected to be paid or received from previous time buckets as there are no assets and liabilities in previous time bucket.
b. We have 5 assets and 5 liabilites.

3: Constraints :
For First Time Bucket:

$$X_1^{1,3} + X_2^{1,4} + X_3^{1,5} + X_4^{1,5} + X_5^{1,6} + X_1 = Y_1^{1,2} + Y_2^{1,3} + Y_3^{1,5} + Y_4^{1,6} \tag{3.3}$$

For Second Time Bucket:

$$[Y_1^{1,2}+Y_1^{1,2}*IA_1^{1,2}*TA_1^{1,2}]+X_1^{2,4}+X_2^{2,4}+X_3^{2,5}+X_4^{2,6}+X_5^{2,7}+G_2 = Y_1^{2,3}+Y_2^{2,4}+Y_3^{2,5}+Y_4^{2,7}+Y_5^{2,7} \tag{3.4}$$

and so on until we have time buckets left.

4: Convert constraint inequality to equality by adding slack variables to the constraints.

3.3

5: Use objective function and constraints of LPP to create initial Simplex table.
6: Find out initial solution by assigning 0 to decision variable
7: Optimality Test:
    a. Calculate c_j - z_j
    b. If the calculated values are positive then optimal solution is the current basic solution. The greatest value column is the key column.
    c. If any one value is negative then choose the greatest value corresponding variable.
8: Feasibility Test:
    Computr the ratios by dividing the value under XB Column by corresponding valueo of key column. The minimum value of ratio identifies the key row.
9: Key Element:
    The intersection of key column and key roe gives the key element.
10: Updating the Table: a. For key row use formula

$$New\_Value = \frac{Old\_Value}{Key\_Value} \tag{3.5}$$

    b. For Other rows use formula:
    New_Value = Old_Value - $\frac{Corresponding\_key\_column\_value*corresponding\_key\_row\_value}{Key\_Value}$
11: Repeat step 7 to 10 until all the values of c_j - z_j are 0 or negative.

3.4 LPP Simplex Method for managing Assets and Liabilities

Example:

$$Max\ Z = 12x_1 + 16x_2$$

Subject to

$$10x_1 + 20x_2 \leq 120$$

$$8x_1 + 8x_2 \leq 80$$

$$x_1, x_2 \geq 0$$

3.5 Solution:

Step 1: Convert all inequality constraints into equalities by the use of slack variables

Let S_1 and S_2 be the two slack variables.

Introducing these slack variables into the inequality constraints and rewriting the objective function such that all variables are on the left hand side of the equation. Model can be rewritten as:

$$Z - 12x_1 - 16x_2 = 0$$

Subject to constraints :

$$10x_1 + 20x_2 + S_1 = 120$$

$$8x_1 + 8x_2 + S_2 = 80$$

$$x_1 , x_2, S_1, S_2 \geq 0$$

3.6

Step 2: Find the initial Basic Feasible Solution. One feasible solution that satisfies all the constraints is:

$x_1 = 0$, $x_2 = 0$, $S_1 = 120$, $S_2 = 80$ and Z = 0

Now S_1, S_2 are Basic Variables

3.7

Step 3: Setup an initial table as:

Row No.	Basic Variable	Coefficient of					Sol.	Ratio
		Z	x_1	x_2	S_1	S_2		
A1	Z	1	-12	**-16**	0	0	0	
B1	S_1	**0**	**10**	**20**	**1**	**0**	**120**	$\frac{120}{20} = 6$
C1	S_2	0	8	**8**	0	1	80	$\frac{80}{8} = 10$

3.8

Step 4: a) Choose the most negative number from row A1. Therefore x_2 is **entering variable**.

b) Calculate Ratio = $\frac{Sol.\ Col}{x_1\ Col\ (x_1 > 0)}$

c) Choose minimum Ratio. That variable S_1 is **departing variable**.

3.9

Step 5: x_2 becomes basic variable and S_1 becomes non basic variable. New table is:

Row No.	Basic Variable	Coefficient of					Sol.	Ratio
		Z	x_1	x_2	S_1	S_2		
A1	Z	1	**-6**	0	0	0	0	
B1	x_2	0	**1/2**	1	1/20	0	6	$\frac{6}{1/2} = 12$
C1	S_2	**0**	**4**	**0**	**-2/5**	**1**	**32**	$\frac{32}{4} = 8$

$$New\ value\ =\ Old\ value\ -\ \frac{Old\ value * Corres.\ key\ row\ value}{Key\ element}$$

3.10 Next Table is:

Row No.	Basic Variable	Coefficient of					Sol.	Ratio
		Z	x_1	x_2	S_1	S_2		
A1	Z	1	0	0	0	0	0	
B1	x_2	0	0	1	1/20	0	2	
C1	x_1	0	1	0	-1/10	$\frac{11}{200}$	8	

Optimal Solution : x_1 = 8, x_2 = 2, Z = 128

3.11 Problem

A	B	C	D	E	F
A3	0			Six month (L3)	5000
A4	0			One year deposit (L4)	15000
A5	0			Three year deposit (L5)	25000
A6	0			Five year deposit (L6)	20000
A7	0			**Total Liability**	**100000**
A8	0				
A9	0				
				Assets	**Interest Rates for Assets**
Objective				Cash	0
				Reserve againse deposit	1.725
Maximize Z	0			Provision	1.725
				Deposit head office	4.5
Constraints				Three months loan	4.65
		InEquality		Six month loan	4.65
1	0	>=	600	One year loan	4.875
2	0	<=	1500	Three year loan	4.9
3	0	>=	600	Five year loan	5.025
4	0	>=	5000		
5	0	>=	5000		
6	0	<=	7500		
7	0	<=	54000		
8	0	>=	0		
9	0	<=	45000		
10	0	<=	60000		
11	0	<=	65000		
12	0	<=	80000		
13	0	<=	20000		
14	0	>=	100000		
15	0	>=	0		
16	0	>=	0		
17	0	>=	0		
18	0	>=	0		
19	0	>=	0		
20	0	>=	0		
21	0	>=	0		
22	0	>=	0		
23	0	>=	0		

3.11 Solution

A	B	C	D	E	F
Variables	**Asset**			**Liability**	
A1	1250			Current deposit (L1)	20000
A2	87500			Three month deposit (L2)	15000
A3	3750			Six month (L3)	5000
A4	0			One year deposit (L4)	15000
A5	0			Three year deposit (L5)	25000
A6	0			Five year deposit (L6)	20000
A7	0			**Total Liability**	**100000**
A8	3750				
A9	3750				
				Assets	**Interest Rates for Assets**
Objective				Cash	0
				Reserve againse deposit	1.725
Maximize Z	1948.125			Provision	1.725
				Deposit head office	4.5
Constraints				Three months loan	4.65
		InEquality		Six month loan	4.65
1	1250	>=	600	One year loan	4.875
2	1250	<=	1500	Three year loan	4.9
3	87500	>=	600	Five year loan	5.025
4	5000	>=	5000		
5	5000	>=	5000		
6	7500	<=	7500		
7	11250	<=	54000		
8	0	>=	0		
9	7500	<=	45000		
10	0	<=	52500		
11	0	<=	57500		
12	0	<=	72500		
13	0	<=	16250		
14	100000	>=	100000		
15	1250	>=	0		
16	87500	>=	0		
17	3750	>=	0		
18	0	>=	0		
19	0	>=	0		
20	0	>=	0		
21	0	>=	0		
22	3750	>=	0		
23	3750	>=	0		

Chapter 4

LSTM implementation for Prediction of Stock Price

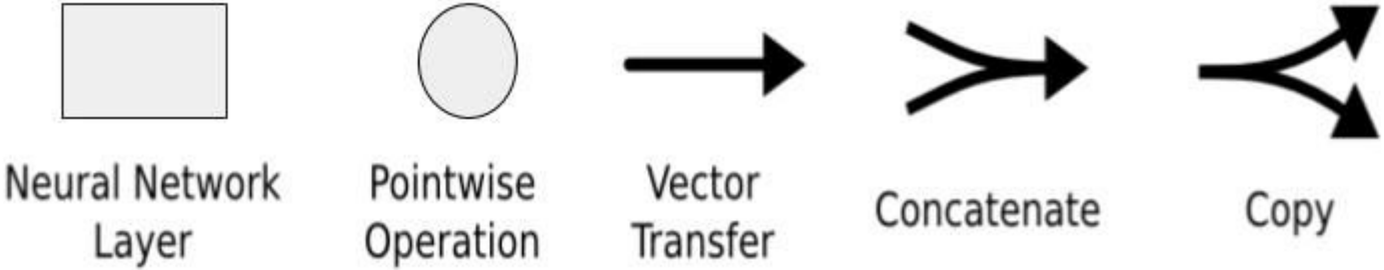
4.1 LSTM implementation for Prediction of Stock Price

- ❖ What is Prediction of Stock Price
- ❖ What is LSTM
- ❖ Why LSTM

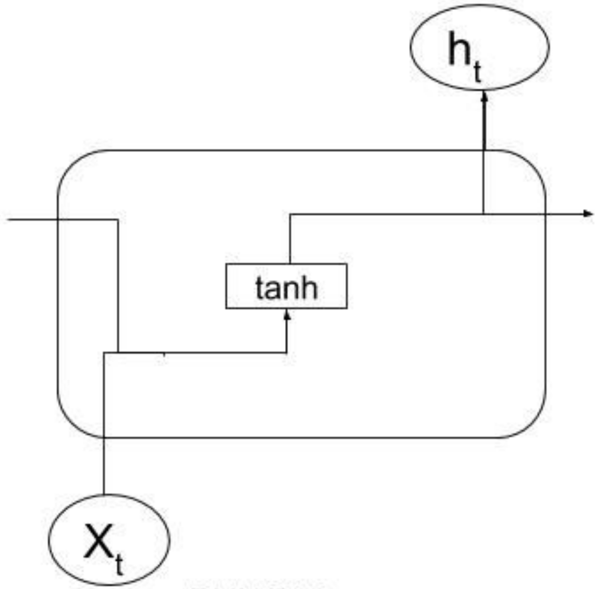
Ex:

“I grew up in Spain… I speak fluent Spanish.”

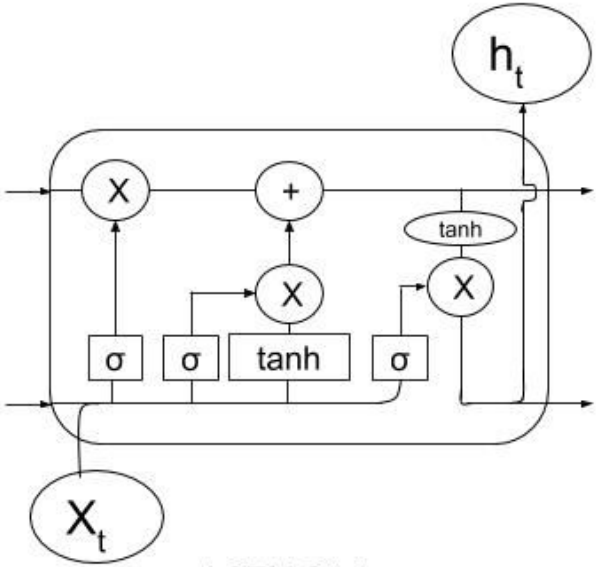
4.2 Notations



4.3 Comparison of RNN and LSTM State

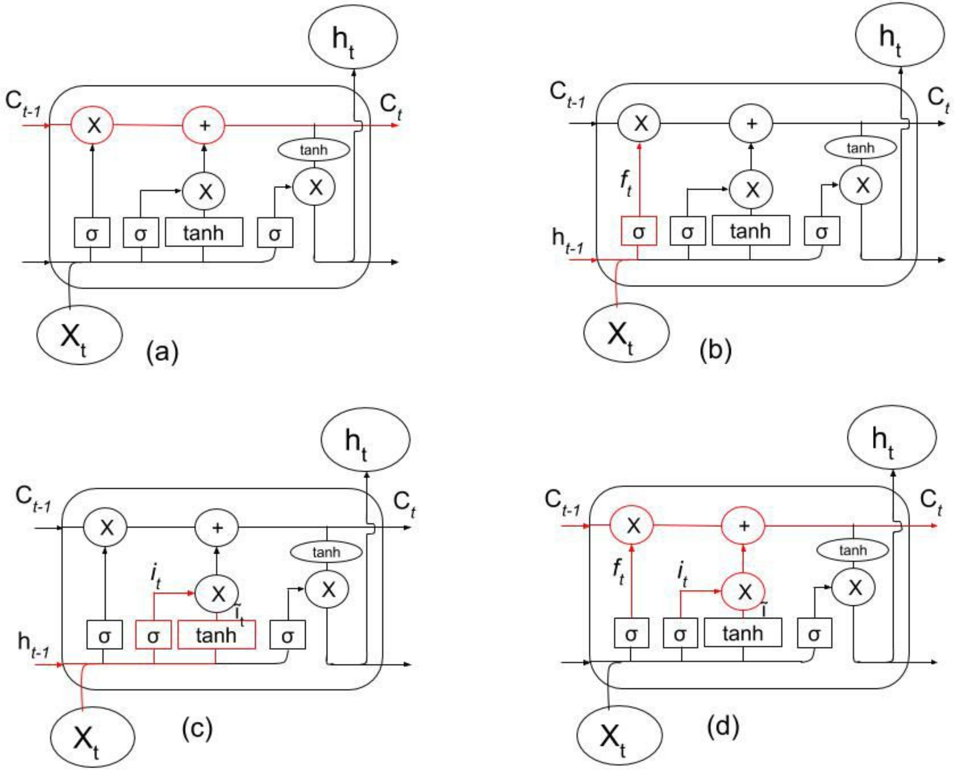


RNN State



LSTM State

4.4 Flow of input through various gates. (a) Cell State, (b) Forget gate, (c) Input Gate and (d) Update to New Cell State



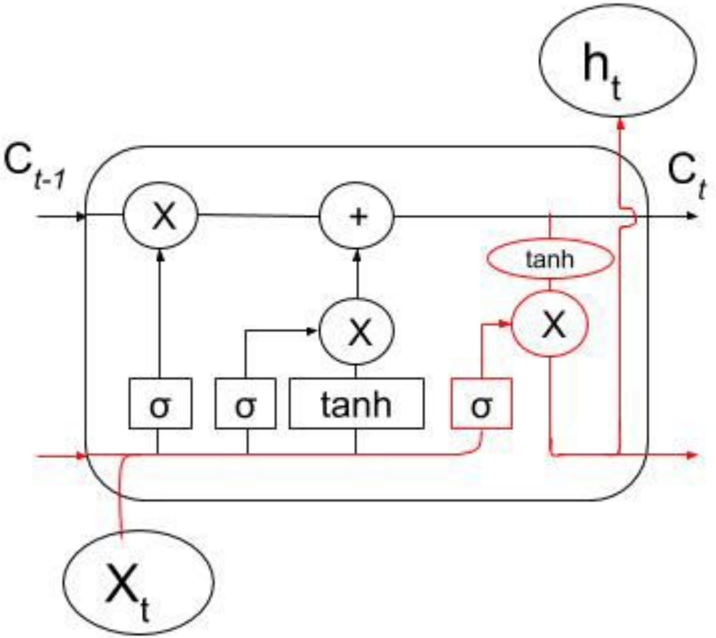
$$f_t = \sigma(W_f.[h_{t-1}, X_t] + b_f)$$

$$i_t = \sigma(W_i.[h_{t-1}, X_t] + b_i)$$

$$\tilde{I}_t = tanh(W_c.[h_{t-1}, X_t] + b_c)$$

$$C_t = f_t * C_{t-1} + i_t * \tilde{I}_t$$

4.5 Output Gate



$$o_t = \sigma(W_o.[h_{t-1}, X_t] + b_o)$$

$$h_t = o_t * tanh(C_t)$$

4.6 Problem Definition:

- **Task:** Prediction of Stock Price
- **Data:** $\{X_i : Open_i, High_i, Low_i\ ;\ Y_i : Close_i\}_{i=1}^{N}$
- **Model:** LSTM:

$$h_t, C_t = LSTM(h_{t-1}, C_{t-1}, X_t)$$

- **Parameters:** W_f, b_f, W_i, b_i, W_c, b_c, W_o, b_o
- **Loss Function:**

$$MeanSquaredError(MSE) = \frac{1}{n}\Sigma_{i=1}^{n}(Predicted\_Y_i - Y_i)^2$$

- **Algorithm:** Adam is the algorithm used for learning the parameters for the LSTM model.

4.7 Computational Complexity:

Computational complexity of LSTM is O(Z),

where $Z = 4 * \#IP * \#h + 4 * \#h^2 + 3 * \#h + \#h * \#OP$,

#IP: Number of inputs

#h: Number of hidden layers

#OP: Number of outputs

4.8 Algorithms to Compare With

- ARIMA
  - AR (Auto Regressive)

$$X_t = c + \sum_{i=1}^{p} \varphi_i X_{t-i} + \varepsilon_t$$

  - I (Integrated)

$$\begin{aligned} y_t^* &= y'_t - y'_{t-1} \\ &= (y_t - y_{t-1}) - (y_{t-1} - y_{t-2}) \\ &= y_t - 2y_{t-1} + y_{t-2} \end{aligned}$$

  - MA (Moving Average)

$$X_t = \mu + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \cdots + \theta_q \varepsilon_{t-q}$$

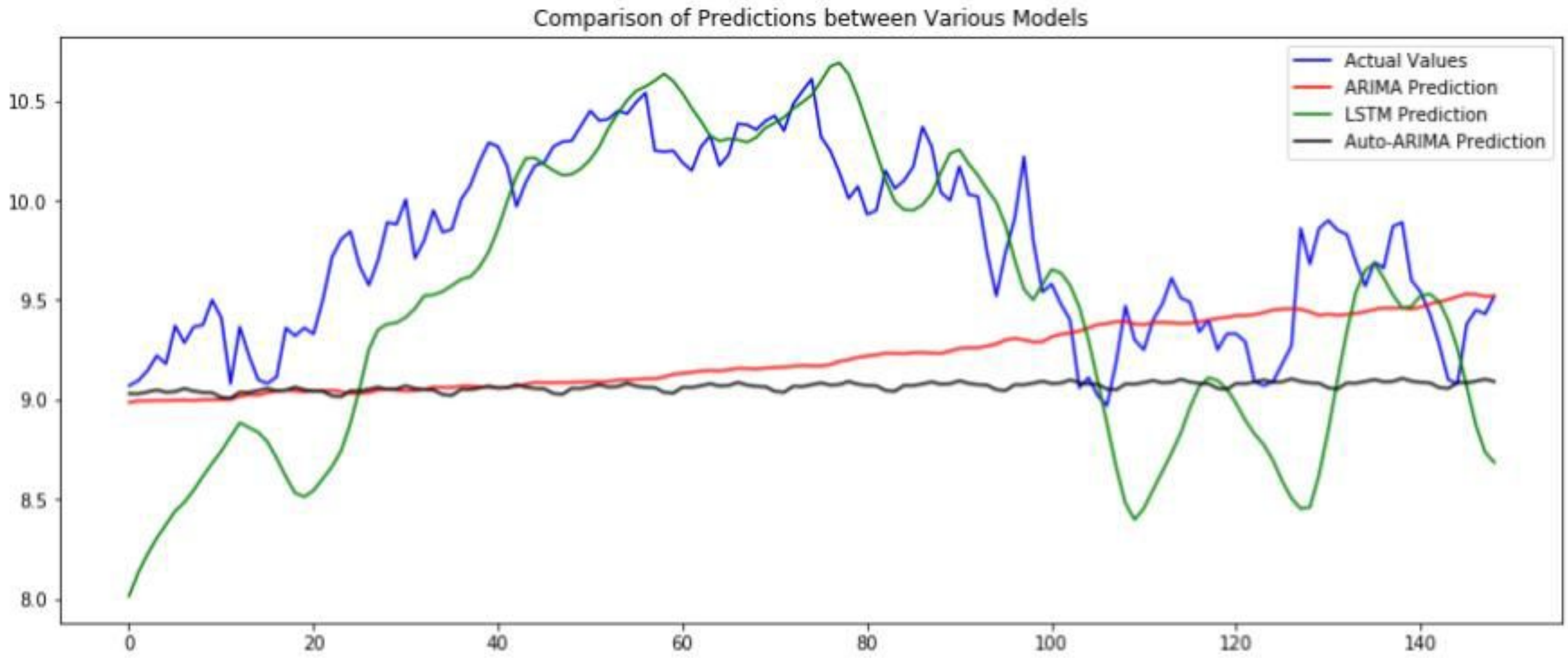
- Auto-ARIMA
  - Select above parameters Automatically

p (lag order): number of lag observations included in the model. Determined using: partial correlation coefficients between the series and lags of itself.

d (degree of differencing): number of times that the raw observations are differenced.
Determined using: ADF Test

q (order of moving average): size of the moving average window
Determined using: AutoCorrelation Function between a time series and lags of itself

4.9 Comparison of Prediction between ARIMA, Auto-ARIMA and LSTM Model with Actual Values {Contribution}



4.10 Results

- MSE on Train Data of LSTM Model: 0.00128 MSE
- MSE on Test Data of LSTM Model: 0.00230 MSE
- MSE on Test Data of ARIMA Model: 0.576 MSE
- MSE on Test Data of Auto-ARIMA Model: 0.695 MSE

Chapter 5

Conclusion and Future Work

5.1 Conclusion

Until now, we have worked on the Liquidity Risk problem and given the following solutions,

1. Simulation of the ALM environment of the bank and gives the brief understanding of the concept
2. Single objective optimization method for liquidity risk management using Linear Programming Formulation
3. An attempt to predict the stock price of the particular stock based on its historical values using LSTM

5.2 Future Work

1. Multiobjective optimization, Nonlinear optimization and Fuzzy control formulation
2. LSTM model can be replaced by attention based models, dual attention based models and also with Multi input based models
3. Feature Engineering using Statistical Approaches

References

[1] Teng, Fan, and Qingwei Zhang. "Empirical Analysis on Interest-Rate Risk to Chinese Life Insurers and Scenario Testing." In 2011 International Conference on Management and Service Science, pp. 1-3. IEEE, 2011.

[2] Vaswani, Ashish, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser, and Illia Polosukhin. "Attention is all you need." In Advances in neural information processing systems, pp. 5998-6008. 2017.

[3] Gülpinar, Nalan, and Dessislava Pachamanova. "A robust optimization approach to asset-liability management under time-varying investment opportunities." Journal of Banking & Finance 37, no. 6 (2013): 2031-2041.

[4] Li, Hongxi, and Guotai Chi. "Assets and Liabilities Portfolio Optimal Model based on ES Controlled Interest Rate Risk." In Proceedings of the International Conference on Business and Information Management, pp. 1-5. ACM, 2017.

[5] Images are drawn in a software package called Jupyter Notebook.

[6] Qin, Yao, Dongjin Song, Haifeng Chen, Wei Cheng, Guofei Jiang, and Garrison Cottrell. "A dual-stage attention-based recurrent neural network for time series prediction." arXiv preprint arXiv:1704.02971 (2017).

[7] Liu, Jian, Yubo Chen, Kang Liu, and Jun Zhao. "Attention-Based Event Relevance Model for Stock Price Movement Prediction." In China Conference on Knowledge Graph and Semantic Computing, pp. 37-49. Springer, Singapore, 2017.

[8] Li, Hao, Yanyan Shen, and Yanmin Zhu. "Stock Price Prediction Using Attention-based Multi-Input LSTM." In Asian Conference on Machine Learning, pp. 454-469. 2018.

[9] Liu, Huicheng. "Leveraging Financial News for Stock Trend Prediction with Attention-Based Recurrent Neural Network." arXiv preprint arXiv:1811.06173 (2018).

[10] https://colab.research.google.com (A cloud based data science work space similar to jupyter notebook).

[11] https://www.rbi.org.in/scripts/NoticationUser.aspx?Id=16&Mode=0

[12] Images are drawn in a software package called RStudio and Jupyter Notebook.

[13] Goodfellow, Ian, Yoshua Bengio, and Aaron Courville. Deep learning. MIT press, 2016.

[14] https://www.tensorflow.org/ (Helpful documentation for learning tensorow, a software library used to design, develop and train deep learning models.)

[15] https://keras.io/ (For creating deep learning model on top of tensoeow.)

[16] Mounika, P.; Sastry, V.N., Cash-ow optimization models for asset liability management (Mtech project report, IDRBT), 2011.

[17] Chaudhury, Rahul; Sastry, V.N., Fuzzy optimization based asset liability management (M-Tech project report, IDRBT), 2014.

[18] Shumway, Robert H., and David S. Stoer. Time series analysis and its applications: with R examples. Springer, 2017.

[19] https://data.world/lpetrocelli/czech-financial-dataset-real-anonymized-transactions/workspace/data-dictionary (Source of dataset used in Chapter 2).

Thank you