Balance Sheet Driven Probability Factorization for Inferring Bank Holdings

[Extended Abstract]

Shawn Mankad
Cornell University
spm263@cornell.edu

Celso Brunetti
The Federal Reserve Board
celso.brunetti@frb.gov

Jeffrey Harris
American University
jharris@american.edu

ACM Reference format:
Shawn Mankad, Celso Brunetti, and Jeffrey Harris. 2017. Balance Sheet Driven Probability Factorization for Inferring Bank Holdings. In *Proceedings of DSMM'17, Chicago, IL, USA, May 14, 2017,* 2 pages.
DOI: http://dx.doi.org/10.1145/3077240.3077243

1 INTRODUCTION

The recent financial crisis has accentuated the need for effective monitoring, oversight and regulation of financial markets and institutions. In response, governments around the world have created new regulatory frameworks. For instance, in the U.S., enhanced oversight and regulation was introduced through the Dodd-Frank act, and similarly the European Union introduced a number of supervisory bodies (European Banking Authority, European Securities and Markets Authority, etc.). Internationally, the Financial Stability Board has been created with the mandate of promoting international financial stability.

One consequence of launching these new regulatory regimes in the Digital age is a vast and increasing amount of data that is available to regulators on the behavior of market participants. For instance, a main thrust of Dodd-Frank was to create transparency and accountability in markets through a myriad of data recording and reporting requirements that apply to exchanges and their market participants. Yet, in spite of increased reporting requirements, regulators have access only to data directly related to their legal purview. Given the significant number of agencies and fragmented nature of the overall oversight ecosystem, it remains a major challenge for regulators to gain deep insight into the balance sheet of financial institutions. Thus, a common and important problem faced by regulatory bodies around the world is how to integrate financial data streams together in a principled manner that yields new insights, thus realizing the potential of these modern regulatory frameworks with more informative risk monitoring systems.

The aim of this work is to help address this key issue in a setting where regulators have access to two linked data sources: (i) equities (stock movements) and (ii) interbank lending data. Stock movements are of course widely and publicly available, whereas the latter data source would be accessible to central banks, like the Federal Reserve Board or the European Central Bank. By using these two data sources, and by building on existing accounting frameworks [3, 4, 9, 10], we partition the balance sheet of the banking sector to isolate the underlying (and unobserved) portfolios held by banks. This partitioning, when combined with balance sheet identities, implies a non-negative matrix factorization problem [6], which has been extensively studied in other domains. Solving the matrix factorization problem provides estimates of a bank's underlying portfolio along different asset classes (equities, bonds, commodities, etc.), thus giving regulators meaningful information synthesized from data sources that are not typically integrated or studied together in this manner.

2 AN ACCOUNTING FRAMEWORK FOR BANKS

We follow the setup of previous works [3, 4, 9, 10]. Before introducing the accounting framework and factorization problem, we first introduce some notation. Let there be n banks under consideration and X be the vector of interbank debt (the total value of liabilities held by other banks). Π_{ij} is the share of bank i's liabilities held by bank j, W_{ik} is the weight invested in each of the k assets by bank i ($\sum_k W_{ik} = 1$), V_{ik} denotes the market value of bank i's assets, including loans to firms and households as well as k asset classes (equities, bonds, commodities, etc.), E_i indicates the market value of bank i's equity, and D_i is the total value of liabilities of bank i held by non-banks.

Consider a financial system in which banks connect lenders to borrowers as intermediaries, collecting deposits from households and firms and investing the deposits in a portfolio of assets, including loans to the household sector (via mortgages and consumer debt) and firms. The balance sheet for a bank can be partitioned as in Table 1.

Assets	Liabilities
$\sum_k W_{ik} V_{ik}$	e_i
	x_i
$\sum_j x_j \Pi_{ij}$	d_i

Table 1: Representation of the balance sheet of a bank i.

We obtain the balance sheet identity by summing assets and liabilities respectively

$$\sum_j x_j \Pi_{ij} + \sum_k W_{ik} V_{ik} = e_i + x_i + d_i. \tag{1}$$

DSMM'17, Chicago, IL, USA

978-1-4503-5031-0/17/05…$15.00
DOI: http://dx.doi.org/10.1145/3077240.3077243

Written using matrix notation (with capital letters), the balance sheet identity is

$$\Pi X + (W \odot V)u = E + X + D, \tag{2}$$

where u is a vector of ones and $\odot$ is the Schur product (element-wise multiplication).

Thus, one can see that W and V are a function of the equity, interbank market, and household debt, which we ignore since it is roughly constant [9].

$$(W \odot V)u = E + (I - \Pi)X. \tag{3}$$

2.1 Probability Factorization

To estimate W, one could pose an optimization problem based on minimizing the Frobenius norm of the difference between a time-series of the combination of stock returns and interbank lending data (captured in $Z = [Z_1, \ldots, Z_T]$) and the estimated factors

$$\begin{aligned} \min_{W,V} \quad & ||Z - WV||_F^2 \qquad (4)\\ \text{subject to} \quad & W \in \mathbb{R}^{n\times K} \\ & V \in \mathbb{R}^{K\times T} \\ & (W)_{ij} \geq 0 \text{ for all } i,j \\ & \sum_{j=1}^{K} (W)_{ij} = 1 \text{ for all } i. \end{aligned}$$

The estimation approach we present alternates between optimizing with respect to W and V. The algorithm solves for W using a projected gradient descent method that has been effective at balancing cost per iteration and convergence rate for similar problems posed in Nonnegative Matrix Factorization [7, 8]. After solving for W, probability constraints are enforced without changing the overall quality of the solution since V can be rescaled without changing the objective function value. Specifically, we have WD and $D^{-1}V$, where D is a diagonal matrix with positive entries, provides different solutions with identical objective function values.

2.1.1 Solving for V. When holding W fixed, the remaining optimization problem written is exactly the usual least squares problem from linear regression.

Starting the objective function,

$$\begin{aligned} O &= ||Z - WV^T||_2^2 && (5)\\ &= (Z - WV^T)^T(Z - WV^T) && (6)\\ &= Z^TZ - VW^TZ - Z^TWV^T + VW^TWV^T. && (7) \end{aligned}$$

Holding W fixed and differentiating with respect to V yields

$$\frac{\partial O}{\partial V} = -2Z^TW + 2VW^TW. \tag{8}$$

Setting the partial derivative equal to zero and solving for V yields

$$V = Z^TW(W^TW)^{-1}. \tag{9}$$

This is the optimal update for the problem $\min_V ||Z - WV^T||_F^2$.

2.1.2 Solving for W. We now turn our attention to solving for W, holding V fixed.

A standard gradient descent algorithm would start with an initial condition $W^{(0)}$ and constants α_i and iterate

(1) For $i = 1, 2, \ldots$
(2) Set $W^{(i+1)} = W^{(i)} - \alpha_i \Delta_W$,

where the gradient of the objective function with respect to W is

$$\Delta_W = WV^TV - ZV.$$

Due to the substraction, the non-negativity of W cannot be guaranteed. Thus, the basic idea of projected gradient descent is to project elements in W to the feasible region using the projection function, which for our problem is defined as $P(\gamma) = max(0, \gamma)$. The basic algorithm is then

(1) For $i = 1, 2, \ldots$
(2) Set $W^{(i+1)} = P(W^{(i)} - \alpha_i \Delta_W)$.

The constants α_i regulate the step size or amount of change in the estimate at each iteration, and converge to zero with i. However, the exact specification of α_i is a main challenge. If the step size is too small, the algorithm will not converge to a stationary point. If the step size is too large, then too many elements of W will be projected to zero and the quality of the estimate will suffer. To guarantee a sufficient decrease at each iteration and convergence to a stationary point, the "Armijo rule" developed in [1, 2] provides a sufficient condition for a given α_i at each iteration

$$||Z - W^{(i+1)}V^T|| - ||Z - W^{(i)}V^T|| \leq \sigma\langle \Delta_{W^{(i)}}, W^{(i+1)} - W^{(i)}\rangle, \tag{10}$$

where $\sigma \in (0, 1)$ and $\langle \cdot, \cdot \rangle$ is the sum of element wise products of two matrices. Thus, for a given α_i, one calculates $W^{(i+1)}$ and checks whether (10) is satisifed. If the condition is satisfied, then the step size α_i is appropriate to guarantee convergence to a stationary point.

3 EMPIRICAL RESULTS

We use data from e-MID interbank market and public stock returns data to form Z and subsequently estimate W over monthly intervals. Our preliminary results of one-way Granger causality[1] of factorization-based variables to the St. Louis Fed Financial Stress Index [5] is an encouraging result. More work is needed to validate the model before using it to gain insight into the balance sheets of banks at a higher frequency than current disclosures allow, and thus better estimate systemic risk and monitor the financial ecosystem.

REFERENCES

[1] D. P. Bertsekas. On the goldstein-levitin-polyak gradient projection method. *Automatic Control, IEEE Transactions on*, 21(2):174–184, 1976.
[2] D. P. Bertsekas. Nonlinear programming. 1999.
[3] C. Brunetti, J. H. Harris, S. Mankad, and G. Michailidis. Interconnectedness in the Interbank Market. *Board of Governors of the Federal Reserve System Finance and Economics Discussion Series*, Oct. 2015.
[4] M. Elliott, B. Golub, and M. O. Jackson. Financial networks and contagion. *American Economic Review*, 104(10):3115–53, 2014.
[5] K. L. Kliesen, D. C. Smith, et al. Measuring financial market stress. *Economic Synopses*, 2010.
[6] D. D. Lee and H. S. Seung. Learning the parts of objects by non-negative matrix factorization. *Nature*, 401:788–791, 10 1999.
[7] C.-b. Lin. Projected gradient methods for nonnegative matrix factorization. *Neural computation*, 19(10):2756–2779, 2007.
[8] S. Mankad, S. Hu, and A. Gopal. Single Stage Prediction with Online Reviews for Mobile App Development and Management. *ArXiv e-prints*, July 2016.
[9] H. S. Shin. Securitisation and financial stability*. *The Economic Journal*, 119(536):309–332, 2009.
[10] H. S. Shin. Financial intermediation and the post-crisis financial system. 2010.

[1] $F = 7.692; p < 0.01$; number of lags is chosen according to AIC; Due to non-stationarity issues, we use the first difference of all variables.