

Optimal Least Squares Approximation through Trace Representation Approach

Sharad Kumar Mahobey¹, Akanksha Dubey²

^{1,2}Department of Mathematics, Shri Rawatpura Sarkar University, Raipur, Chhattisgarh, India

Abstract - In this paper, an optimal model is proposed for least square approximation (LSA) based on linear approach. This model comprises with the characteristics of Woodbury matrix identity (WMI). WMI is based on square and invertible matrices. The input for WMI is two square and invertible matrices and the output is two corresponding matrices hold the identity. This is referred as WMI. The application of this WMI is presented in this paper as LSA. The proposed WMI-LSA model follows the trace representation of LSA under the WMI linear correspondence. The sketch of this model as follows: Given the linear equation contained the parameters as input and output, there are some unknown parameters also as the coefficients, there is the objective function to be minimized. The existed LSA model provides either efficiency or security but this compact model has dual characteristics as efficiency and security simultaneously. The proposed model has the short algorithm steps for fast application. The small keys and compact formula provide the compact advantage for applying it in modern AI applications. This model has the following application: artificial intelligence (AI), machine learning (ML), medicine etc.

Keywords--LSA, WMI, Matrix, Trace Representation, AI, ML.

I. INTRODUCTION

This paper proposes the optimal least square approximation based on Woodbury matrix identity. This optimality is possible by Matrix-Trace approach. The proposed model provides the security and efficiency both in context to its application. The proposed algorithm is followed compact and error-correcting based approach for fast and errorless performance. Speed by trace representation and efficiency by WMI are the two key novelty and significance of the proposed WMI-Trace Representation based LSA model in comparison with existing AI least-squares and Woodbury Matrix Identity based approaches. These AI based compact matrix schemes achieves the modern security and efficiency standards.

Whenever required any estimation to plan the project for its effective completion as the optimal and probable errorless output, the LSA exists. Artificial intelligence, machine learning, whether forecasting, diagnosis, medicine etc. are some key examples of LSA.

Majorly, there are two classes of such approximations; linear and non-linear. Although both carries the unique advantage, but non-linear approximation follows the non-trivial algorithms but the linear approximation lies with trivial and compact algorithms.

This paper presents the linear LSA models. In 1986, Antoulas et al. [1] studied the interpolation problem based on scalar rational function. The literatures [2, 3, 9, 13, 19, 23] present the foundation of approximation and its discrete applications. In literature [2], the foundation of LSA is given.

The elementary to advance LSA is explained nicely. The literature [3] focuses on the application of LSA. There are several methods exist for applying into LSA. The illustration is also presented in lucid way. The literature [9] approaches the multidimensional mechanism on LSA based generalizations. In literature [13] compact algorithms are illustrated using discrete structures of the existed algorithms.

There are pre-LSA and post-LSA applications given in literature [19]. The literature [13] sets the parameters for generalizing the optimality of LSA. In 2002, Chahlaoui [4] published the results on time reduction model through the minimized linear regression model. Demmel [5] observed the decomposition varieties in structured matrices in 2000.

In 2015, a result on quadrature approximation is published by Drmac et al. [6]. Edelman et al. [7] illustrated the approximation over orthogonality constraints in 1998. Ellacott [8] derived a discrete result on complex plane for generating linear approximation over rational points. Golub et al. [8] presented an innovative linear approximation results based on pseudo-inverse approach on derivatives.

Gustavsen [11, 12] studied the vector fitting for generalizing pole relocation. In 2006, Hendrickx et al. [14] established an innovative vector fitting technique carried the rational frequency domain for non linear approximation. In 2018, Hokanson et al. [15] presented a model as the non linear least square approximation.

Ionit [16] proposed an optimal system for approximating the large scale dynamical system in 2013. Same year, Lefteriu et al. [17] generalized the vector fitting method through its convergence testing.

The problem function is presented in the next section.

Problem Function: Let the linear equation be,

$$(1) \dots Y = a_0 + a_1 X_1 + a_2 X_2 + \dots + a_n X_n,$$

Where, Y is the output, X is the input and $a_0, a_1, \dots, a_n$ are the unknown coefficients.

There are two objectives:

Objective 1: Given: $a_0, a_1, \dots, a_n$.

To estimate: $e_0, e_1, e_2, \dots, e_n$.

Applied function: One-to-one correspondence as follows:

$$\begin{array}{cccc} a_0 & a_1 & & a_n \\ \updownarrow & \updownarrow & . & \updownarrow \\ e_0 & e_1 & & e_n \end{array}$$

Objective 2: Given: X, Y .

To estimate: Y' .

Applied function: The optimal objective function as follows:

$$\min_{a_0, a_1, \dots, a_n} \|Y' - Y\|_2.$$

Where,

$$(2) \dots Y' = e_0 + e_1 X_1 + e_2 X_2 + \dots + e_n X_n.$$

Next, WMI is presented.

WMI: This is a lemma stated as follows:

WMI Lemma: The identity is,

$$(3) \dots (A + UCV)^{-1} = A^{-1} - A^{-1} \cup (C^{-1} + VA^{-1}U)^{-1} \times VA^{-1}.$$

Where,

$$A \in R^{n \times n}, C \in R^{m \times m}, U \in R^{n \times m}, V \in R^{m \times n}.$$

Next, WMI application is presented for LSA.

WMI based LSA: The result of an optimal LSA is obtained over the generalization of WMI presented as follows:

Theorem: WMI corresponds with trace (LSA) representation.

Proof: From WMI lemma,

$$(A + UCV)^{-1} = A^{-1} - A^{-1} \cup (C^{-1} + VA^{-1}U)^{-1} \times VA^{-1}.$$

Where,

$$A = \begin{bmatrix} a_{11} & & & \\ & \ddots & & \\ & & \ddots & \\ & & & a_{nn} \end{bmatrix}$$

$$C = \begin{bmatrix} c_{11} & & & \\ & \ddots & & \\ & & \ddots & \\ & & & c_{mm} \end{bmatrix}$$

$$U = \begin{bmatrix} u_{11} & & & \\ & \ddots & & \\ & & \ddots & \\ & & & u_{nm} \end{bmatrix}$$

$$V = \begin{bmatrix} v_{11} & & & \\ & \ddots & & \\ & & \ddots & \\ & & & v_{mn} \end{bmatrix}$$

The trace representation of the matrices A, C, U, V is presented as follows:

$$Tr(A) = \sum_{i=1}^n a_{ii} = a_{11} + a_{22} + \dots + a_{nn}.$$

$$Tr(C) = \sum_{i=1}^m c_{ii} = c_{11} + c_{22} + \dots + c_{mm}.$$

$$Tr(U) = \sum_{i=1}^n u_{ii} = u_{11} + u_{22} + \dots + u_{nm}.$$

$$Tr(V) = \sum_{i=1}^n v_{ii} = v_{11} + v_{22} + \dots + v_{nn}.$$

These trace repetitions follow the properties of invariance under transpose, linearity, cyclic property and sum of eigenvalues. Beginning with the invariance property,

$$Tr(A) = Tr(A^T).$$

$$Tr \begin{pmatrix} a_{11} & & & \\ & \cdot & & \\ & & \cdot & \\ & & & a_{nn} \end{pmatrix} = Tr \begin{pmatrix} a_{11} & & & \\ & \cdot & & \\ & & \cdot & \\ & & & a_{nn} \end{pmatrix}^T = Tr \begin{pmatrix} a_{11} & & & \\ & \cdot & & \\ & & \cdot & \\ & & & a_{nn} \end{pmatrix}.$$

$$Tr(C) = Tr(C^T).$$

$$Tr \begin{pmatrix} c_{11} & & & \\ & \cdot & & \\ & & \cdot & \\ & & & c_{nn} \end{pmatrix} = Tr \begin{pmatrix} c_{11} & & & \\ & \cdot & & \\ & & \cdot & \\ & & & c_{nn} \end{pmatrix}^T = Tr \begin{pmatrix} c_{11} & & & \\ & \cdot & & \\ & & \cdot & \\ & & & c_{nn} \end{pmatrix}.$$

$$Tr(U) = Tr(U^T).$$

$$Tr \begin{pmatrix} u_{11} & & & \\ & \cdot & & \\ & & \cdot & \\ & & & u_{nn} \end{pmatrix} = Tr \begin{pmatrix} u_{11} & & & \\ & \cdot & & \\ & & \cdot & \\ & & & u_{nn} \end{pmatrix}^T = Tr \begin{pmatrix} v_{11} & & & \\ & \cdot & & \\ & & \cdot & \\ & & & v_{nn} \end{pmatrix}.$$

$$Tr(V) = Tr(V^T).$$

$$Tr \begin{bmatrix} v_{11} & & \\ & \ddots & \\ & & v_{nn} \end{bmatrix} = Tr \begin{bmatrix} v_{11} & & \\ & \ddots & \\ & & v_{nn} \end{bmatrix}^T = Tr \begin{bmatrix} v_{11} & & \\ & \ddots & \\ & & v_{nn} \end{bmatrix}.$$

Linearity:

$$Tr(A + B) = Tr(A) + Tr(B).$$

$$Tr(kA) = kTr(A).$$

$$Tr \begin{bmatrix} a_{11} & & \\ & \ddots & \\ & & a_{nn} \end{bmatrix} + \begin{bmatrix} b_{11} & & \\ & \ddots & \\ & & b_{nn} \end{bmatrix} = Tr \begin{bmatrix} a_{11} & & \\ & \ddots & \\ & & a_{nn} \end{bmatrix} + Tr \begin{bmatrix} b_{11} & & \\ & \ddots & \\ & & b_{nn} \end{bmatrix} \cdot C$$

cyclic:

$$Tr(AB) = Tr(BA).$$

$$Tr \begin{bmatrix} a_{11} & & \\ & \ddots & \\ & & a_{nn} \end{bmatrix} \begin{bmatrix} b_{11} & & \\ & \ddots & \\ & & b_{nn} \end{bmatrix} = Tr \begin{bmatrix} b_{11} & & \\ & \ddots & \\ & & b_{nn} \end{bmatrix} \begin{bmatrix} a_{11} & & \\ & \ddots & \\ & & a_{nn} \end{bmatrix}.$$

Sum of Eigenvalues:

$$Tr(A) = \sum_{i=1}^n \lambda_i = \lambda_1 + \lambda_2 + \dots + \lambda_n.$$

Thus,

$$(A + UCV)^{-1} =$$

$$\left(\begin{bmatrix} a_{11} & & & \\ & \ddots & & \\ & & \ddots & \\ & & & a_{nn} \end{bmatrix} + \begin{bmatrix} u_{11} & & & \\ & \ddots & & \\ & & \ddots & \\ & & & u_{nn} \end{bmatrix} \begin{bmatrix} c_{11} & & & \\ & \ddots & & \\ & & \ddots & \\ & & & c_{mm} \end{bmatrix} \begin{bmatrix} v_{11} & & & \\ & \ddots & & \\ & & \ddots & \\ & & & v_{mn} \end{bmatrix} \right)^{-1}$$

$$A^{-1} - A^{-1} \cup (C^{-1} + VA^{-1}U)^{-1} \times VA^{-1} =$$

$$\begin{bmatrix} a_{11} & & & \\ & \ddots & & \\ & & \ddots & \\ & & & a_{nn} \end{bmatrix}^{-1} -$$

$$\begin{bmatrix} a_{11} & & & \\ & \ddots & & \\ & & \ddots & \\ & & & a_{nn} \end{bmatrix}^{-1} \cup$$

$$\left(\begin{bmatrix} c_{11} & & & \\ & \ddots & & \\ & & \ddots & \\ & & & c_{mm} \end{bmatrix}^{-1} + \begin{bmatrix} v_{11} & & & \\ & \ddots & & \\ & & \ddots & \\ & & & v_{mn} \end{bmatrix} \begin{bmatrix} a_{11} & & & \\ & \ddots & & \\ & & \ddots & \\ & & & a_{nn} \end{bmatrix}^{-1} \begin{bmatrix} u_{11} & & & \\ & \ddots & & \\ & & \ddots & \\ & & & u_{nn} \end{bmatrix} \right)^{-1}$$

$$\times \begin{bmatrix} v_{11} & & & \\ & \ddots & & \\ & & \ddots & \\ & & & v_{mn} \end{bmatrix} \begin{bmatrix} a_{11} & & & \\ & \ddots & & \\ & & \ddots & \\ & & & a_{nn} \end{bmatrix}^{-1}$$

The LSA problem:

$$.Y = a_0 + a_1X_1 + a_2X_2 + \dots + a_nX_n.$$

The linear combination of eigenvalues is represented as the trace of the matrix. Thus, the trace (LSA) is represented as follows:

$$(a_0, a_1, \dots, a_n) \leftrightarrow (\lambda_0, \lambda_1, \dots, \lambda_n).$$

Thus,

$$.Y = \lambda_0 + \lambda_1X_1 + \lambda_2X_2 + \dots + \lambda_nX_n.$$

The following correspondence exists,

$$(A + UCV)^{-1} \leftrightarrow Y.$$

2. Trace Representation:

$$.LSA : Tr(Y) = Tr(a_0 + a_1X_1 + a_2X_2 + \dots + a_nX_n).$$

$$WMI : Tr[(A + UCV)^{-1}] = Tr[A^{-1} - A^{-1} \cup (C^{-1} + VA^{-1}U)^{-1} \times VA^{-1}]$$

3. Processing by theorem (4.1):

$$(A^{-1} - A^{-1} \cup (C^{-1} + VA^{-1}U)^{-1} \times VA^{-1}) \leftrightarrow Y'.$$

4. Output:

$$(A^{-1} - A^{-1} \cup (C^{-1} + VA^{-1}U)^{-1} \times VA^{-1}) \leftrightarrow Y'.$$

This implies that,

$$Y' \leftrightarrow Y.$$

This completes the proof.

Next WMI-LSA model is presented.

WMI-LSA Model: The algorithm as follows:

1. Input:

$$LSA : .Y = a_0 + a_1X_1 + a_2X_2 + \dots + a_nX_n.$$

WMI:

$$(A + UCV)^{-1} = A^{-1} - A^{-1} \cup (C^{-1} + VA^{-1}U)^{-1} \times VA^{-1}.$$

$$.Y' = e_0 + e_1X_1 + e_2X_2 + \dots + e_nX_n.$$

Next, the flow chart of the WMI-LSA model is presented:

Flow Chart of the Proposed WMI-LSA Model:

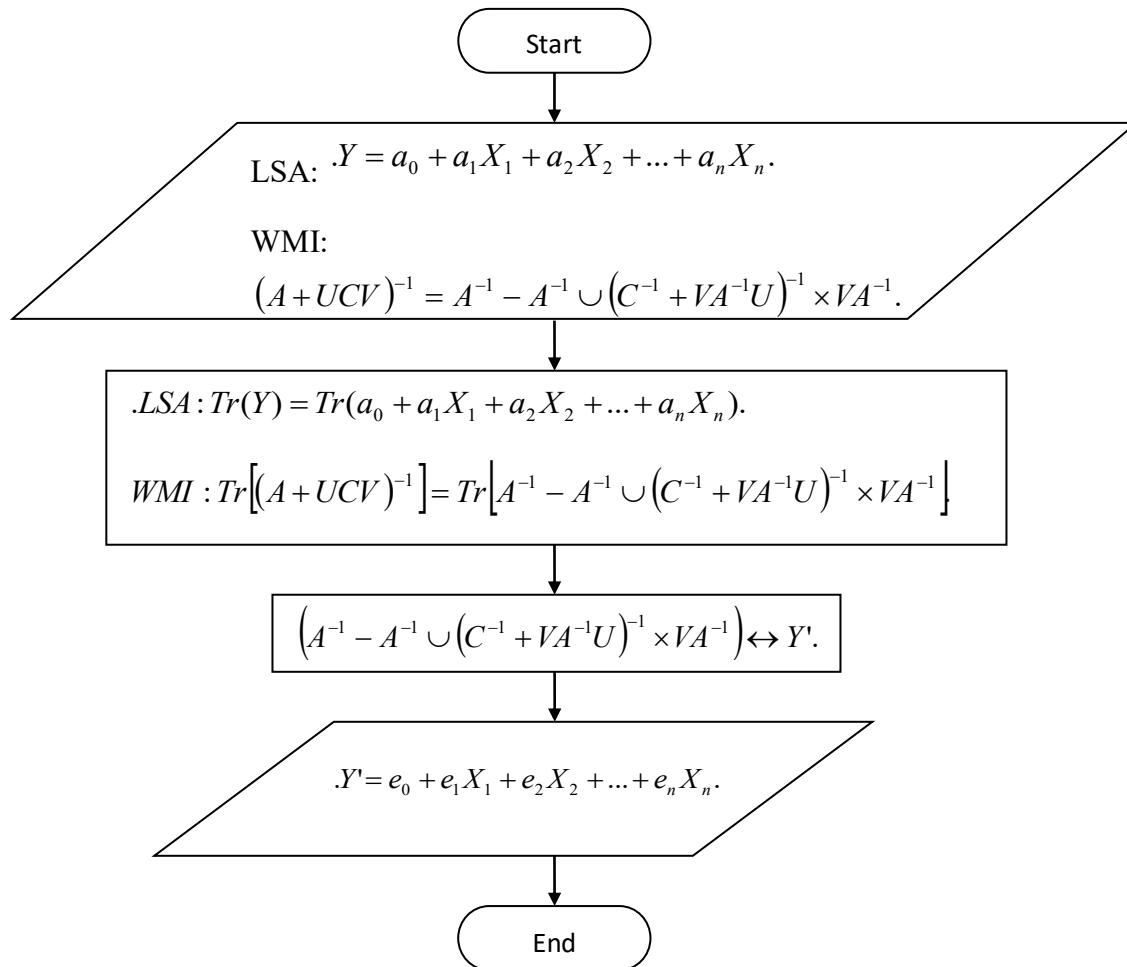


Figure 1. Flow Chart of WMI-LSA Model.

Next, the experiment model is illustrated.

WMI-LSA Proto Model: The proto model is presented with the input and output records as follows:

1. *Input:*

LSA: $Y = 2 + 3X_1$.

$$\text{WMI: } (A + UCV)^{-1} = \left(\begin{bmatrix} 1 & 0 & 1 \\ 0 & 2 & 0 \\ 0 & 1 & 3 \end{bmatrix} + \begin{bmatrix} u_{11} & & \\ & \ddots & \\ & & u_{mm} \end{bmatrix} \begin{bmatrix} c_{11} & & \\ & \ddots & \\ & & c_{mm} \end{bmatrix} \begin{bmatrix} v_{11} & & \\ & \ddots & \\ & & v_{mm} \end{bmatrix} \right)^{-1}.$$

2. Trace Representation:

$$\text{LSA: } \text{Tr}(Y) = \sum_{i=1}^n y_{ii}.$$

$$\text{WMI: } \text{Tr}[(A + UCV)^{-1}] = \sum_{i=1}^n A_{ii}.$$

3. Processing by theorem (4.1):

$$\sum_{i=1}^n y_{ii} \leftrightarrow \sum_{i=1}^n A_{ii}.$$

4. Output:

$$Y' = \begin{bmatrix} e_{11} & & \\ & \ddots & \\ & & e_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}.$$

The computational complexity of the proposed model is presented in next section.

Computational Complexity: WMI-LSA model counts the point through the following vectors:

$$\text{Tr}(A) = \sum_{i=1}^n a_{ii} = a_{11} + a_{22} + \dots + a_{nn}.$$

$$\text{Tr}(C) = \sum_{i=1}^n c_{ii} = c_{11} + c_{22} + \dots + c_{mm}.$$

$$\text{Tr}(U) = \sum_{i=1}^n u_{ii} = u_{11} + u_{22} + \dots + u_{mm}.$$

$$\text{Tr}(V) = \sum_{i=1}^n v_{ii} = v_{11} + v_{22} + \dots + v_{mm}.$$

Each vector has the dimension for generating the security towards its computation over the matrix representation as follows:

$$Tr(A) \leftrightarrow \begin{bmatrix} a_{11} & & \\ & \ddots & \\ & & a_{nn} \end{bmatrix}.$$

$$Tr(C) \leftrightarrow \begin{bmatrix} c_{11} & & \\ & \ddots & \\ & & c_{nn} \end{bmatrix}.$$

$$Tr(U) \leftrightarrow \begin{bmatrix} u_{11} & & \\ & \ddots & \\ & & u_{nn} \end{bmatrix}.$$

$$Tr(V) \leftrightarrow \begin{bmatrix} v_{11} & & \\ & \ddots & \\ & & v_{nn} \end{bmatrix}.$$

Thus, the 1-1 correspondence is defined as follows:

$$Tr(A) \leftrightarrow Tr(C) \leftrightarrow Tr(U) \leftrightarrow Tr(V).$$

$$\Rightarrow \begin{bmatrix} a_{11} & & \\ & \ddots & \\ & & a_{nn} \end{bmatrix} \leftrightarrow \begin{bmatrix} c_{11} & & \\ & \ddots & \\ & & c_{nn} \end{bmatrix} \leftrightarrow \begin{bmatrix} u_{11} & & \\ & \ddots & \\ & & u_{nn} \end{bmatrix} \leftrightarrow \begin{bmatrix} v_{11} & & \\ & \ddots & \\ & & v_{nn} \end{bmatrix}.$$

The comparison between the proposed model and existed model is presented in next.

Comparison between WMI-LSA Model and Existed Modern Model: The first comparison with minimization objective function based models, viz. OLS, WLS, AM etc. These models implies the keys over the closed form solution presented by,

$$\hat{\beta} = (X^T X)^{-1} (X^T Y).$$

This CFS based approach is not sufficient for designing the secure model due to its trivial invertible conditions. The proposed model comprises with the WMI-trace mechanism for establishing smaller keys with 512-bit probable security under the following LSA-WMI doors:

$$.LSA : Tr(Y) = Tr(a_0 + a_1 X_1 + a_2 X_2 + \dots + a_n X_n).$$

$$WMI : Tr[(A + UCV)^{-1}] = Tr[A^{-1} - A^{-1} \cup (C^{-1} + VA^{-1}U)^{-1} \times VA^{-1}]$$

II. CONCLUSION

The proposed WMI-LSA model is optimal and compact in the parameter of trace and linearity respectively. Any LSA can be transformed into an optimal WMI-LSA model through this model. The proposed LSA model has efficiency and security both characteristics simultaneously.

The proposed model has the short algorithm steps for fast application. The small keys and compact formula provide the compact advantage for applying it in modern AI applications. The theorem on WSI-LSA correspondence sets the trace representation of problem statement which provides the eigenvalues for generating the equivalent objective function.

This minimized objective function is possible by the trace of the input linear function, where the sum of the diagonal elements of the matrices of the linear equations. There are several applications exist over this optimal WMI-LSA model. There are several possibilities exist; AI-LSA, ML-LSA, AI-WMI, AI-WMI-LSA etc. The proposed fast algorithm sets the compact application in context to formulation and generalization both. Thus, the proposed dual WMI-LSA approach sets the compact and optimal approximation.

Declarations

Funding Source (Optional): No Funding received.

Acknowledgement (Optional): We acknowledge to Department of Mathematics, Shri Rawatpura Sarkar University, Raipur, Chhattisgarh, India.

Data Availability (Optional): Available on request.

Authors' Contribution (Optional): This paper is created by both authors simultaneously.

**Use of AI and AI-Assisted Technologies:* No AI tool is used.

**Conflict of Interest:* No conflict of interest.

**Copyright Permissions:* Yes.

REFERENCES

- [1] A. C. Antoulas and B. D. Q. Anderson, On the scalar rational interpolation problem, IMA J. Math. Control I., 3 (1986), pp. 61–88, <https://doi.org/10.1093/imamci/3.2-3.61>.
- [2] G. A. J. Baker and P. Graves-Morris, Pad'e Approximants, Cambridge, second ed., 1996.
- [3] J.-P. Berrut and L. N. Trefethen, Barycentric Lagrange interpolation, SIAM Rev., 46 (2004), pp. 501–517.
- [4] Y. Chahlaoui and P. V. Dooren, A collection of benchmark examples for model reduction of linear time invariant dynamical systems, Tech. Report 2, SLICOT, Feb. 2002.
- [5] J. Demmel, Accurate singular value decompositions of structured matrices, SIAM J. Matrix Anal. Appl., 21 (2000), pp. 562–580.
- [6] Z. Drmac, S. Gugercin, and C. Beattie, Quadrature-based vector fitting for discretized H2 approximation, SIAM J. Sci. Comput., 37 (2015), pp. A625–A652.
- [7] A. Edelman, T. A. Arias, and S. T. Smith, The geometry of algorithms with orthogonality constraints, SIAM J. Matrix Anal. Appl., 20 (1998), pp. 303 – 353.
- [8] S. Ellacott and J. Williams, Rational Chebyshev approximation in the complex plane, SIAM J. Numer. Anal., 13 (1976), pp. 310–323.
- [9] A. Gil, J. Segura, and N. M. Temme, Numerical Methods for Special Functions, SIAM, 2007.
- [10] G. H. Golub and V. Pereyra, The differentiation of pseudo-inverses and nonlinear least squares problems whose variables separate, SIAM J. Numer. Anal., 10 (1973), pp. 413–432, <https://doi.org/10.1137/0710036>.
- [11] B. Gustavsen, Improving the pole relocating properties of vector fitting, IEEE Trans. Power Deliver., 21 (2006), pp. 1587–1592.
- [12] B. Gustavsen and A. Semlyen, Rational approximation of frequency domain responses by vector fitting, IEEE Trans. Power Deliver., 14 (1999), pp. 1052–1061.
- [13] P. C. Hansen, V. Pereyra, and G. Scherer, Least Squares Data Fitting with Applications, Johns Hopkins University Press, 2013.
- [14] W. Hendrickx and T. Dhaene, A discussion of “rational approximation of frequency domain responses by vector fitting”, IEEE Trans. Power Deliver., 21 (2006), pp. 441–443.
- [15] J. M. Hokanson and C. C. Magruder, Projected nonlinear least squares for H2 model reduction, 2018, <https://arxiv.org/abs/XXXXXXX>.
- [16] A. C. Ionit, a, Lagrange rational interpolation and its application to large-scale dynamical systems, PhD thesis, Rice University, 2013.
- [17] S. Lefteriu and A. C. Antoulas, On the convergence of the vector-fitting algorithm, IEEE T. Microw. Theory, 61 (2013), pp. 1435–1443.
- [18] Y. Nakatsukasa, O. S'ete, and L. N. Trefethen, The AAA algorithm for rational approximation, 2016, <https://arxiv.org/abs/1612.00337v2>.
- [19] J. Nocedal and S. J. Wright, Numerical Optimization, Springer, New York, 2006. 20 JEFFREY M. HOKANSON AND CALEB C. MAGRUDER. Nakatsukasa et al. [18] presented an AAA algorithm based on the counting the rational points.
- [20] C. K. Sanathanan and J. Koerner, Transfer function synthesis as a ratio of two complex polynomials, IEEE T. Automat. Contr., 8 (1963), pp. 56–58. Sanathanan et al. [20] published the original results on transfer function. This transfer function applies on the two complex polynomials. In 2010, Schreier et al. [21] presented a creative analogue of signal processing. This mechanism carries the complex valued data. In 2016, a new vector fitting algorithm is developed by Shi.
- [21] P. J. Schreier and L. L. Scharf, Statistical Signal Processing of Complex-Valued Data: Theory of Improper and Noncircular Signals, Cambridge University Press, Cambridge, 2010.
- [22] G. Shi, On the nonconvergence of the vector fitting algorithm, IEEE T. Circuits-II, 63 (2016), pp. 718–722.
- [23] L. N. Trefethen, Approximation Theory and Approximation Practice, SIAM, 2013.