# STUDIES ON EFFICIENT ALGORITHMS FOR CONVEX AND NON-CONVEX PROBLEMS WITH APPLICATIONS IN SIGNAL PROCESSING, COMMUNICATION, AND MACHINE LEARNING

**R. Jyothi**

**Centre for Applied Research in Electronics**
**Indian Institute of Technology Delhi**
**March, 2022**

# STUDIES ON EFFICIENT ALGORITHMS FOR CONVEX AND NON-CONVEX PROBLEMS WITH APPLICATIONS IN SIGNAL PROCESSING, COMMUNICATION, AND MACHINE LEARNING

**by**


**R. Jyothi**

**Centre for Applied Research in Electronics**


Submitted

*in fulfillment of the requirements of the degree of Doctor of Philosophy*

to the

**Indian Institute of Technology Delhi**

**March, 2022**

*Dedicated to*

***My Parents***

# Certificate

This is to certify that the thesis entitled **"Studies on efficient algorithms for convex and non-convex problems with applications in signal processing, communication, and machine learning"** being submitted by **Ms. R. Jyothi** to the Centre for Applied Research in Electronics (CARE), Indian Institute of Technology Delhi, for the award of the degree of **Doctor of Philosophy** is the record of the bonafide research work carried out by her under our supervision. In our opinion, the thesis has reached the standards fulfilling the requirements of the regulations relating to the degree.

The results contained in this thesis have not been submitted either in part or in full to any other university or institute for the award of any degree or diploma.


(Prof. Prabhu Babu)
CARE
Indian Institute of Technology Delhi
Hauz Khas, New Delhi 110016
India

(Prof. Rajendar Bahl)
CARE
Indian Institute of Technology Delhi
Hauz Khas, New Delhi 110016
India

# Acknowledgements

First and foremost, I would like to thank my guide, Prof. Prabhu Babu, for having faith in me and giving me an opportunity to work with him. I am extremely grateful to him for introducing me to the field of optimization. His thoughtful advice has helped me a lot over the last five years of study. I am also thankful to Prof. Rajendar Bahl, Prof. Arun Kumar, Prof. Ananjan Basu, and Prof. Saif Khan Mohammed for their suggestions during my work. I would like to thank all my friends at Signal Processing, CARE for supporting me during my study. I am also thankful to all my other IIT Delhi friends for making my stay at Delhi enjoyable. Especially, I would like to thank my best friend Rijul, who has always offered to help at any time. Finally, I would like to thank my mom, dad, and brother for their unconditional love and moral support.

R. Jyothi

R. Jyothi

# Abstract

The efficiency of an optimization algorithm plays a crucial role in the performance of several engineering applications. In this thesis, the focus is on utilizing the Majorization Minimization (MM) framework to develop efficient iterative optimization algorithms. In particular, MM-based algorithms are proposed to solve convex, non-convex, and non-smooth problems that have applications in the field of signal processing, communication, and machine learning. Also, for some problems, the combination of the MM framework with other commonly used optimization algorithms, such as the Alternating Direction Method of Multipliers (ADMM), is explored. One of the motivations behind studying the MM framework is its flexibility in solving a broad spectrum of problems. For instance, apart from solving convex problems, it can solve problems whose cost functions are non-convex and non-smooth. Also, it can handle problems with non-convex constraints. Another key advantage of MM is that it helps in exploiting the problem structure to develop problem-specific optimization algorithms.

In the first part of the thesis, the focus is on developing efficient optimization algorithms for the problems encountered in the field of signal processing. In this regard, four different problems related to signal processing applications are considered - 1) Source localization, 2) Non-negative matrix factorization, 3) Atomic Norm Minimization and, 4) Toeplitz covariance matrix estimation problems. For the first problem, which is a non-convex and non-smooth problem, an optimization algorithm named SOLVIT is proposed. SOLVIT is an MM-based iterative algorithm which, unlike the state-of-the-art algorithms, tackles the problem exactly without any approximations. Next, the non-negative matrix factorization problem, which is a non-convex problem

with multiple optimization variables, is studied. By inspecting the structure of this problem, two different iterative optimization algorithms named INOM and PARINOM are presented. While the former optimization algorithm is based on the block MM framework - an extension of the MM technique employed to solve problems with several block variables, the latter optimization algorithm is based on the principle of the MM procedure. Subsequently, for the convex Atomic Norm Minimization and the non-convex Toeplitz covariance matrix estimation problems, iterative optimization algorithms named DYANOM and ATOM are proposed. To develop both the optimization algorithms, a hidden structure of the underlying optimization problem is exploited. Moreover, DYANOM is based on the integration of the MM framework with Dykstra's projection. Also, two versions of ATOM, based on the combination of MM with ADMM and MM with Dykstra's projection, are proposed. For each problem, numerical simulations are conducted, which indicate that the proposed optimization algorithms have superior performance when compared to the state-of-the-art algorithms.

The second part of the thesis deals with tackling two different optimization problems that have applications in the field of communication. The first problem is concerned with the construction of unimodular sequences with good autocorrelation properties. To attain such sequences, an iterative optimization algorithm named SLOPE is proposed that minimizes the maximum peak sidelobe level of a unimodular sequence. Therefore, SLOPE optimizes a non-convex minimax problem. SLOPE arrives at a stationary point of the problem by leveraging the adaptation of the MM framework for minimax problems coupled with the utilization of the Mirror Descent Algorithm (MDA). Later, SLOPE is extended to handle additional constraints such as the energy, peak-to-average-power ratio, and spectral constraints. Numerical simulations confirm that SLOPE can generate sequences of considerably longer lengths with lower peak sidelobe levels. Next, a different non-convex minimax problem is optimized to construct optimal frames. Such frames comprises of a set of unit norm vectors such that the maximum correlation among them is minimal. To solve the minimax problem, two iterative optimization algorithms named TELET and FLIP with different updating schemes are proposed. While the former MM-based algorithm updates all the columns comprising the frame vectors at once, the latter block MM-based algorithm updates the frame in a column-wise fashion. In addition to utilizing the MM framework, TELET and FLIP borrow MDA and state-of-the-art linear programming optimization algorithms, respectively, to construct optimal frames. Numerical simulations reveal that TELET and FLIP can construct frames of longer lengths with lower coherence values.

In the third part of the thesis, the focus is on solving optimizations problems that have applications in the field of machine learning. Under this field, the maximum likelihood parameter estimation for the classical multinomial logistic regression classifier is considered. Although the problem is convex and differentiable, it does not enjoy a closed-form solution. To solve this problem, an MM-based iterative algorithm named PIANO is proposed, which can parallely update each element of the optimization variable. Moreover, unlike the state-of-the-art algorithms, PIANO can also be easily modified to handle the parameter estimation problem for the sparse multinomial logistic regression classifier. Simulations are conducted to compare PIANO with the existing methods, and it is found that PIANO performs better than the existing methods in terms of speed of convergence. Then, the $\ell_p$ norm linear regression optimization problem, which has applications in semi-supervised learning, is studied. To tackle this problem, an iterative optimization algorithm named PROMPT is proposed based on the MM technique. Unlike the state-of-the-art algorithms, PROMPT can handle the underlying cost function for any value of $p$. Moreover, PROMPT updates each element of the optimization variable parallelly. Numerical simulations highlight that PROMPT performs better than the state-of-the-art algorithms in terms of speed of convergence.

# सारांश

एक अनुकूलन एल्गोरिथ्म की दक्षता कई इंजीनियरिंग अनुप्रयोगों के प्रदर्शन में महत्वपूर्ण भूमिका निभाती है। इस थीसिस में, कुशल पुनरावृत्त अनुकूलन एल्गोरिदम विकसित करने के लिए मेजराइजेशन मिनिमाइजेशन (एमएम) ढांचे का उपयोग करने पर ध्यान केंद्रित किया गया है। विशेष रूप से, एमएम-आधारित एल्गोरिदम उत्तल, गैर-उत्तल, और . को हल करने के लिए प्रस्तावित हैं गैर-चिकनी समस्याएं जिनमें सिग्नल प्रोसेसिंग, संचार और मशीन सीखने के क्षेत्र में अनुप्रयोग हैं। इसके अलावा, कुछ समस्याओं के लिए, एमएम फ्रेमवर्क का संयोजन अन्य सामान्य रूप से उपयोग किए जाने वाले अनुकूलन एल्गोरिदम के साथ, जैसे कि अल्टरनेटिंग गुणक की दिशा विधि (ADMM) की खोज की गई है। एमएम ढांचे का अध्ययन करने के पीछे की प्रेरणाओं में से एक समस्याओं के व्यापक स्पेक्ट्रम को हल करने में इसकी लचीलापन है। उदाहरण के लिए, उत्तल समस्याओं को हल करने के अलावा, यह उन समस्याओं को भी हल कर सकता है जिनकी लागत कार्य गैर-उत्तल और गैर-चिकनी हैं। साथ ही, यह गैर-उत्तल बाधाओं के साथ समस्याओं को संभाल सकता है। MM का एक अन्य प्रमुख लाभ यह है कि यह समस्या-विशिष्ट अनुकूलन एल्गोरिदम विकसित करने के लिए समस्या संरचना का दोहन करने में मदद करता है।

थीसिस के पहले भाग में, सिग्नल प्रोसेसिंग के क्षेत्र में आने वाली समस्याओं के लिए कुशल अनुकूलन एल्गोरिदम विकसित करने पर ध्यान केंद्रित किया गया है। इस संबंध में, सिग्नल प्रोसेसिंग अनुप्रयोगों से संबंधित चार अलग-अलग समस्याओं पर विचार किया जाता है - 1) स्रोत स्थानीयकरण, 2) गैर-नकारात्मक मैट्रिक्स कारककरण, 3) परमाणु मानदंड न्यूनतमकरण, और, 4) Toeplitz सहप्रसरण मैट्रिक्स अनुमान समस्याएं। बाकी समस्या के लिए, जो एक गैर-उत्तल और गैर-चिकनी समस्या है, SOLVIT नामक एक अनुकूलन एल्गोरिथ्म प्रस्तावित है। SOLVIT एक MM-आधारित पुनरावृत्त एल्गोरिथ्म है, जो अत्याधुनिक एल्गोरिदम के विपरीत, बिना किसी अनुमान के समस्या से निपटता है। इसके बाद, गैर-ऋणात्मक मैट्रिक्स गुणनखंडन समस्या का अध्ययन किया जाता है, जो कई अनुकूलन चरों के साथ एक गैर-उत्तल समस्या है। इस समस्या की संरचना का निरीक्षण करके, INOM और PARINOM नामक दो अलग-अलग पुनरावृत्ति अनुकूलन एल्गोरिदम प्रस्तुत किए जाते हैं। जबकि पूर्व अनुकूलन एल्गोरिथ्म ब्लॉक MM ढांचे पर आधारित है - कई ब्लॉक चर के साथ समस्याओं को हल करने के लिए नियोजित MM तकनीक का एक विस्तार, बाद वाला अनुकूलन एल्गोरिथ्म MM प्रक्रिया के सिद्धांत पर आधारित है। इसके बाद, उत्तल परमाणु मानक न्यूनीकरण और गैर-उत्तल टोप्लिट्ज सहप्रसरण मैट्रिक्स आकलन समस्याओं के लिए, पुनरावृत्त अनुकूलन एल्गोरिदम DYANOM और ATOM नाम प्रस्तावित हैं। दोनों अनुकूलन एल्गोरिदम विकसित करने के लिए, अंतर्निहित अनुकूलन समस्या की एक छिपी संरचना का शोषण किया जाता है। इसके अलावा, DYANOM डायक्स्ट्रा के प्रक्षेपण के साथ MM ढांचे के एकीकरण पर आधारित है। साथ ही, ADMM के साथ MM के संयोजन के आधार पर ATOM के दो संस्करण और डायक्स्ट्रा के प्रक्षेपण के साथ एमएम प्रस्तावित हैं। प्रत्येक समस्या के लिए, संख्यात्मक सिमुलेशन आयोजित किए जाते हैं, जो इंगित करते हैं कि प्रस्तावित अनुकूलन एल्गोरिदम का अत्याधुनिक एल्गोरिदम की तुलना में बेहतर प्रदर्शन है।

थीसिस का दूसरा भाग दो अलग-अलग अनुकूलन समस्याओं से निपटने से संबंधित है जिनका संचार के क्षेत्र में अनुप्रयोग है। पहली समस्या अच्छे ऑटोसहसंबंध गुणों के साथ अनिमोड्यूलर अनुक्रमों के निर्माण से संबंधित है। ऐसे अनुक्रमों को प्राप्त करने के लिए, SLOPE नामक एक पुनरावृत्त अनुकूलन एल्गोरिथ्म प्रस्तावित है जो एक अनिमोड्यूलर अनुक्रम के अधिकतम शिखर साइडलोब स्तर को कम करता है। इसलिए, SLOPE एक गैर-उत्तल न्यूनतम अधिकतम समस्या का अनुकूलन करता है। मिरर डिसेंट एल्गोरिथम के उपयोग के साथ मिलकर मिनिमैक्स समस्याओं के लिए एमएम फ्रेमवर्क के अनुकूलन का लाभ उठाकर स्लोप समस्या के एक स्थिर बिंदु पर आता है (एमडीए)। बाद में, SLOPE को अतिरिक्त बाधाओं जैसे कि ऊर्जा, शिखर-से-औसत-शक्ति अनुपात, और वर्णक्रमीय बाधाओं को संभालने के लिए बढ़ाया गया है। संख्यात्मक सिमुलेशन इस बात की पुष्टि करते हैं कि SLOPE निचली

चोटी के साथ काफी लंबी लंबाई के अनुक्रम उत्पन्न कर सकता है साइडलोब स्तर। अगला, एक अलग गैर-उत्तल मिनीमैक्स समस्या को इष्टतम फ्रेम बनाने के लिए अनुकूलित किया गया है। इस तरह के फ्रेम में यूनिट नॉर्म वैक्टर का एक सेट होता है, जिसमें उनके बीच अधिकतम सहसंबंध न्यूनतम होता है। मिनिमैक्स समस्या को हल करने के लिए, अलग-अलग अद्यतन योजनाओं के साथ TELET और FLIP नामक दो पुनरावृत्त अनुकूलन एल्गोरिदम प्रस्तावित हैं। जबकि पूर्व MM-आधारित एल्गोरिथम फ्रेम वैक्टर वाले सभी कॉलम को एक साथ अपडेट करता है, बाद वाला ब्लॉक MM-आधारित एल्गोरिथम कॉलम-वार फैशन में फ्रेम को अपडेट करता है। MM ढांचे का उपयोग करने के अलावा, TELET और FLIP इष्टतम फ्रेम बनाने के लिए क्रमशः एमडीए और अत्याधुनिक रैखिक प्रोग्रामिंग अनुकूलन एल्गोरिदम उधार लेते हैं। संख्यात्मक सिमुलेशन से पता चलता है कि TELET और FLIP कम सुसंगतता मूल्यों के साथ लंबी लंबाई के फ्रेम का निर्माण कर सकते हैं।

थीसिस के तीसरे भाग में, उन अनुकूलन समस्याओं को हल करने पर ध्यान केंद्रित किया गया है जिनमें मशीन सीखने के पुराने समय में अनुप्रयोग हैं। इस क्षेत्र के तहत, क्लासिकल मल्टीनोमियल लॉजिस्टिक रिग्रेशन क्लासिफायरियर के लिए अधिकतम संभावना पैरामीटर अनुमान पर विचार किया जाता है। हालांकि समस्या उत्तल और अलग-अलग है, यह नहीं है एक बंद फॉर्म समाधान का आनंद लें। इस समस्या को हल करने के लिए, पियानो नामक एक एमएम-आधारित पुनरावृत्त एल्गोरिदम प्रस्तावित है, जो अनुकूलन चर के प्रत्येक तत्व को समानांतर रूप से अद्यतन कर सकता है। इसके अलावा, अत्याधुनिक एल्गोरिदम के विपरीत, विरल बहुपद के लिए पैरामीटर अनुमान समस्या को संभालने के लिए पियानो को आसानी से संशोधित किया जा सकता है लॉजिस्टिक रिग्रेशन क्लासी एर। मौजूदा तरीकों के साथ पियानो की तुलना करने के लिए सिमुलेशन आयोजित किए जाते हैं, और यह पाया जाता है कि पियानो अभिसरण की गति के मामले में मौजूदा तरीकों से बेहतर प्रदर्शन करता है। फिर, एलपी मानदंड रैखिक प्रतिगमन अनुकूलन समस्या, जिसमें अर्ध-पर्यवेक्षित शिक्षण में अनुप्रयोग हैं, का अध्ययन किया जाता है। इससे निपटने के लिए समस्या, MM तकनीक के आधार पर PROMPT नामक एक पुनरावृत्त अनुकूलन एल्गोरिथ्म प्रस्तावित है। अत्याधुनिक एल्गोरिदम के विपरीत, PROMPT p के किसी भी मूल्य के लिए अंतर्निहित लागत फ़ंक्शन को संभाल सकता है। इसके अलावा, PROMPT अनुकूलन चर के प्रत्येक तत्व को समानांतर रूप से अद्यतन करता है। संख्यात्मक सिमुलेशन इस बात पर प्रकाश डालते हैं कि PROMPT अभिसरण की गति के मामले में अत्याधुनिक एल्गोरिदम से बेहतर प्रदर्शन करता है।

# Contents

# List of Figures

# List of Tables

# Abbreviations

| | |
|---|---|
| ADMM | Alternating Direction Method of Multipliers |
| ANM | Atomic Norm Minimization |
| APG | Accelerated Proximal Gradient |
| ATOM | Alternating projection based TOeplitz covariance Matrix |
| BCD | Block Coordinate Descent |
| BSS | Blind Source Separation |
| CAN | Cyclic Algorithm New |
| CDMA | Code Division Multiple Access |
| CE | Circulant Embedding |
| CRLB | Cramer-Rao Lower bound |
| CS | Compressed Sensing |
| DoA | Direction of Arrival |
| DYANOM | DYkstra's projection based Atomic NOrM solver |
| EM | Expectation Maximization |
| ETF | Equiangular Tight Frames |
| FA | Factor Analysis |
| FFT | Fast Fourier Transform |
| FLIP | Frame design using LInear Programming |
| GF | Grassmannian Frame |
| HALS | Hierarchical Alternating Least Square |
| INOM | Iterative NOnnegative Matrix Factorization |
| IRLS | Iterative Reweighted Least Squares |
| ISL | Integrated Sidelobe Level |

| | |
|---|---|
| ISPM | Incoherence via Shifted Power Method |
| KKT | Karush Kuhn Tucker |
| KL | Kullback-Leibler |
| LP | Linear Programming |
| MDA | Mirror Descent Algorithm |
| MIMO | Multiple Input Multiple Output |
| MM | Majorization Minimization |
| MLE | Maximum Likelihood Estimation |
| MLR | Multinomial Logistic Regression |
| MU | Multiplicative Update |
| NMF | Non-negative Matrix Factorization |
| PAPR | Peak-to-Average-Power Ratio |
| PARINOM | Parallel Iterative NOnnegative Matrix Factorization |
| PCA | Principal Component Analysis |
| PIANO | Parallel Iterative Algorithm for MultiNomial LOgistic Regression classifier |
| PROMPT | Parallel iteRative algOrithM for $\ell_P$ norm regression via MajorizaTion Minimization |
| PSD | Positive Semi-Definite |
| PSL | Peak Sidelobe Level |
| POCS | Projection onto Convex Sets |
| RD-LS | Range-Differences by the Least-Square |
| R-LS | Range Least Squares |
| RMSE | Root Mean Square Error |
| RIR | Room Impulse Response |
| SCM | Sample Covariance Matrix |
| SDP | SemiDenite Programming |
| SIDCO | Sequential Iterative Decorrelation by Convex Optimization |
| SFP | Standard Fixed Point |
| SLOPE | Sequence with LOw Peak sidelobE |
| SNR | Signal to Noise ratio |
| SOLVIT | SOurce Localization Via an ITerative technique |

| | |
|---|---|
| SVD | Singular Value Decomposition |
| SQUAREM | Squared Iterative method |
| TDOA | Time Difference of Arrival |
| TELET | TEchnique to devise Large dimensional Equiangular Tight frames |
| TOA | Time of Arrival |
| TSC | Toeplitz Structured Covariance |
| UNTF | Unit Norm Tight Frame |

# Notations

| | |
|---|---|
| a | Scalar |
| $\boldsymbol{a}$ | Vector |
| $\mathbf{A}$ | Matrix |
| $\mathbb{R}$,$\mathbb{C}$ | Set of real and complex numbers, respectively |
| $\mathbb{R}^{m\times 1}$, $\mathbb{C}^{m\times 1}$ | Set of $m$ dimensional vectors of real and complex numbers, respectively |
| $\mathbb{R}^{m\times m}$, $\mathbb{C}^{m\times m}$ | Set of $m \times m$ matrices of real numbers and complex numbers, respectively |
| $\mathbb{H}^{m\times m}$ | Set of $m \times m$ Hermitian matrices |
| $\mathbf{R}^{+}$ | Set of non-negative real numbers. |
| $(\cdot)^T$ | Transpose |
| $(\cdot)^*$ | Complex conjugate |
| $(\cdot)^H$ | Complex conjugate transpose |
| $(\cdot)^{-1}$ | Inverse |
| $\odot$ | Hadamard product |
| $\otimes$ | Kronecker product |
| $\lvert\mathbf{A}\rvert$ | Determinant of matrix $\mathbf{A}$ |
| $\mathrm{Re}(\cdot)$ | Real part |
| $\mathrm{Tr}(\mathbf{A})$ | Trace of matrix $\mathbf{A}$ |
| $\mathrm{Rank}(\mathbf{A})$ | Rank of matrix $\mathbf{A}$. |
| $\mathrm{Vec}(\mathbf{A})$ | Vector obtained by stacking the columns of $\mathbf{A}$ |
| $\mathbf{I}_N$ | Identity matrix of size $N$ |
| $\nabla f$ | Gradient of a function $f$ |
| $\nabla^2 f$ | Hessian of a function $f$. |
| $\mathbf{A} \succeq 0$ | $\mathbf{A}$ is a PSD matrix |

| | |
|---|---|
| $\mathbf{A} \succ 0$ | $\mathbf{A}$ is a positive definite matrix |
| $\mathbf{A} - \boldsymbol{B} \succeq 0$ | Matrix $\mathbf{A} - \boldsymbol{B}$ is PSD |
| $\mathbf{A} - \boldsymbol{B} \succ 0$ | Matrix $\mathbf{A} - \boldsymbol{B}$ is positive definite |
| $\lambda_{max}(\mathbf{A})$ | Maximum eigenvalue of matrix $\mathbf{A}$ |
| $\lambda_{min}(\mathbf{A})$ | Minimum eigenvalue of matrix $\mathbf{A}$ |
| $\lvert\boldsymbol{a}\rvert$ | Element-wise absolute of the vector $\boldsymbol{a}$ |
| $\lVert\boldsymbol{a}\rVert_2$ | Euclidean norm of the vector $\boldsymbol{a}$ |
| $\lVert\boldsymbol{a}\rVert_1$ | $\ell_1$ norm of the vector $\boldsymbol{a}$ |
| $\lVert\boldsymbol{a}\rVert_p$ | $\ell_p$ norm of the vector $\boldsymbol{a}$ |
| $\lVert\mathbf{A}\rVert_F$ | Frobenious norm of matrix $\mathbf{A}$ |
| $\mathbf{E}[\cdot]$ | Statistical expectation |
| $\max(\mathbf{A}, \boldsymbol{B})$ | Element wise maximum between $\mathbf{A}$ and $\boldsymbol{B}$. |
| $\circ$ | Element wise multiplication. |
| $\oslash$ | Element wise division. |