RGC Ref. No.:
UGC/FDS14/P03/20
(please insert ref. above)

RESEARCH GRANTS COUNCIL
COMPETITIVE RESEARCH FUNDING SCHEMES FOR THE LOCAL SELF-FINANCING DEGREE SECTOR

FACULTY DEVELOPMENT SCHEME (FDS)

Completion Report

(for completed projects only)

Submission Deadlines:
1. *Auditor's report with unspent balance, if any: within **six** months of the approved project completion date.*
2. *Completion report: within **12** months of the approved project completion date.*

Part A: The Project and Investigator(s)

1. Project Title

Low-rank Matrix Optimization via Nonconvex Regularization with Applications

2. Investigator(s) and Academic Department(s) / Unit(s) Involved

Research Team	Name / Post	Unit / Department / Institution
Principal Investigator	YU Kwok-wai / Associate Professor	Department of Mathematics, Statistics and Insurance / The Hang Seng University of Hong Kong
Co-Investigator(s)	YANG Xiao Qi / Professor	Department of Applied Mathematics / The Hong Kong Polytechnic University
Co-Investigator(s)	CHOY Siu Kai / Professor	Department of Mathematics, Statistics and Insurance / The Hang Seng University of Hong Kong

3. Project Duration

	Original	Revised	Date of RGC / Institution Approval *(must be quoted)*
Project Start Date	1 July 2021		
Project Completion Date	30 June 2023	31 December 2023	22 November 2022
Duration *(in month)*	24 months	30 months	22 November 2022
Deadline for Submission of Completion Report	30 June 2024	31 December 2024	22 November 2022

4.4 Please attach photo(s) of acknowledgement of RGC-funded facilities / equipment.
NA

Part B: The Final Report

5. Project Objectives

5.1 Objectives as per original application

1. To investigate exact recovery property and the recovery bounds of the nonconvex S_p regularization for low-rank matrix optimization in terms of model error, absolute deviation and ℓ_2 consistency under the restricted eigenvalue condition, and to derive the second-order growth property of the S_p regularized function at local minima.

2. To propose a proximal gradient method with continuation technique (PGMC) to solve the S_p regularization problem and investigate the convergence theory of the algorithms to the ground true low-rank solution under a weak regularity condition.

3. To conduct numerical experiments to show the numerical capability of the PGMC on solving the S_p regularization problem and compare with several state-of-the-art algorithms for low-rank matrix optimization. To apply our S_p regularization model and numerical algorithm to complete scRNA-seq data and use the enhanced data to carry out downstream analyses including the clustering of cell population, cell type determination and gene regulatory network inference.

5.2 Revised objectives

Date of approval from the RGC:	NA
Reasons for the change:	NA

5.3 Realisation of the objectives
(Maximum 1 page; please state how and to what extent the project objectives have been achieved; give reasons for under-achievements and outline attempts to overcome problems, if any)

We have addressed three key objectives in this project.

The first objective was to investigate exact recovery property and the recovery bounds of the nonconvex S_p regularization for low-rank matrix optimization, focusing on model error, absolute deviation and ℓ_2 consistency under the restricted eigenvalue condition. To achieve this, we explored various nonconvex lower-order regularization methods for solving low-rank matrix optimization problems and extended the concept of ℓ_p norm to our proposed regularization method. Additionally, we conducted a theoretical study of the

Sp regularization problem. By assuming the restricted eigenvalue condition (REC), we examined the exact recovery property and recovery bound of the regularization problem.

For the second objective, we proposed a proximal gradient method with continuation technique (PGMC) for solving the *Sp* regularization problem. The primary computational effort of the PGMC involves the proximal optimization subproblem, which has a closed-form solution for certain values of p and can be efficiently solved using Newton's method in more general cases. To fulfill this objective, we reviewed existing convergence theorems and refined them to clearly state that the PGMC converges at a linear rate toward an approximation of the ground true solution, contingent upon the additional REC assumption. We also investigated the convergence theory of our algorithms to the ground true low-rank solution under a weak regularity condition.

Finally, for the third objective, we conducted numerical experiments to demonstrate the effectiveness of the PGMC in solving the S_p regularization problem, comparing it with several state-of-the-art algorithms for low-rank matrix optimization. To achieve this, we generated random simulation data using standard process of low-rank matrix optimization and applied various advanced algorithms to analyze this data. Our numerical experiments validated the convergence results and rates of the PGMC toward the ground true low-rank matrix, highlighting the robust low-rank promoting capability of the S_p regularization through phase diagram studies. Furthermore, we applied our proposed S_p regularization model and numerical algorithm to impute single-cell RNA sequencing (scRNA-seq) data, utilizing the enhanced data for downstream analyses. These analyses not only assessed the accuracy of the imputed gene expression profile but also provided significant biological insights.

5.4 Summary of objectives addressed to date

Objectives *(as per 5.1/5.2 above)*	**Addressed** *(please tick)*	**Percentage Achieved** *(please estimate)*
1. To investigate exact recovery property and the recovery bounds of the nonconvex S_p regularization for low-rank matrix optimization in terms of model error, absolute deviation and ℓ_2 consistency under the restricted eigenvalue condition, and to derive the second-order growth property of the S_p regularized function at local minima.	✓	100%
2. To propose a proximal gradient method with continuation technique (PGMC) to solve the S_p regularization problem and investigate the convergence theory of the algorithms to the ground true low-rank solution under a weak regularity condition.	✓	100%
3. To conduct numerical experiments to show the numerical capability of the PGMC on solving the S_p regularization problem and compare with several state-of-the-art algorithms for low-rank matrix optimization. To apply our S_p regularization model and numerical algorithm to complete scRNA-seq data and use the enhanced data to carry out	✓	100%

downstream analyses including the clustering of cell population, cell type determination and gene regulatory network inference.		

6. Research Outcome

6.1 Major findings and research outcome
(Maximum 1 page; please make reference to Part C where necessary)

In this project, we explore low-rank matrix optimization problems through nonconvex regularization. We establish theoretical foundations and develop algorithms based on the nonconvex S_p regularization to tackle low-rank matrix optimization problems. Our proposed proximal gradient method with continuation technique (PGMC) is specifically designed to address the S_p regularization problem. A key finding from our work is that the PGMC demonstrates efficiency in numerical experiments, including analyses single-cell RNA sequencing data. The numerical results indicate that our proposed method effectively solves low-rank matrix optimization problems. Additionally, we investigate the exact recovery property and recovery bounds of the S_p regularization model, and we establish the convergence theory for the PGMC.

6.2 Potential for further development of the research and the proposed course of action *(Maximum half a page)*

This project focuses on low-rank matrix optimization problems. We conduct numerical experiments, specifically analyzing single-cell RNA sequencing data. The proposed proximal gradient method with continuation technique (PGMC) is employed to efficiently address these problems. We have established theoretical foundations and developed algorithms based on nonconvex S_p regularization for low-rank matrix optimization. These theories and methods are not only valuable but also adaptable to a wide range of fields in science and engineering, including system identification and control, Euclidean embedding, collaborative filtering, machine learning, and computer vision.

7. Layman's Summary

(Describe in layman's language the nature, significance and value of the research project, in no more than 200 words)

The explosive growth of big data presents significant opportunities and has greatly benefited the development and the transformation of various disciplines. However, the challenges arise from issues such as missing data and gross errors, which complicate big data analysis. Low-rank matrix optimization is a crucial technique for addressing these challenges and has been successfully applied across various fields. Among the methods available, nuclear norm regularization is one of the most popular and practical approaches for solving low-rank matrix optimization problems. Nevertheless, it has certain limitations in both theoretical properties and practical applications. consequently, there is a pressing need to develop alternative low-rank promoting techniques that offer strong theoretical property and robust numerical performance.

In this project, we propose a proximal gradient method with continuation technique to tackle low-rank matrix optimization problems through nonconvex regularization. We also investigate its convergence theory toward the true low-rank solution. Our project enhances the understanding of nonconvex lower-order regularization problems from both theoretical and algorithmic perspectives, as well as their applications to a wide range of practical issues.

Part C: Research Output

8. Peer-Reviewed Journal Publication(s) Arising Directly From This Research Project

(Please attach a copy of the publication and/or the letter of acceptance if not yet submitted in the previous progress report(s). All listed publications must acknowledge RGC's funding support by quoting the specific grant reference.)

The Latest Status of Publications									
Year of Publication	**Year of Acceptance** *(For paper accepted but not yet published)*	**Under Review**	**Under Preparation** *(optional)*	**Author(s)** *(denote the correspond-ing author with an asterisk*)*	**Title and Journal / Book** *(with the volume, pages and other necessary publishing details specified)*	**Submitted to RGC** *(indicate the year ending of the relevant progress report)*	**Attached to this Report** *(Yes or No)*	**Acknowledged the Support of RGC** *(Yes or No)*	**Accessible from the Institutional Repository** *(Yes or No)*
2022				Yaohua Hu, Gang Li*, Minghua Li, and Carisa Kwok Wai Yu	Multiple-sets split quasi-convex feasibility problems: adaptive subgradient methods with convergence guarantee, *Journal of Nonlinear Variational Analysis*, 6(2), pp. 15-33.	**2022**	**Yes**	**Yes (Annex I)**	**No** https://doi.org/10.23952/jnva.6.2022.2.03
2024				Yaohua Hu, Jingchao Li, Yanyan Liu, and Carisa Kwok Wai Yu*	Quasi-subgradient Methods with Bregman Distance for Quasi-convex Feasibility Problems, *Journal of Nonlinear and Variational Analysis*, 8(3), pp. 381-395.	**N/A**	**Yes**	**Yes (Annex II)**	**Yes** https://researchdb.hsu.edu.hk/view/publication/202400154
	2024			Yaohua Hu, Xinlin Hu, Carisa Kwok Wai Yu, and Jing Qin*	Joint Sparse Optimization: Lower-order Regularization Method and Application in Cell Fate Conversion, *Inverse Problems*, accepted.	**N/A**	**No**	**Yes**	**No**
		2024		Yin Li, Carisa Kwok Wai Yu, Jinhua Wang, and Weiping Shen	Convergence Rate of the Projection Gradient Algorithm for Split Quality Problems in Hilbert Space, *Journal of Nonlinear and Convex Analysis*, under review.	**N/A**	**No**	**Yes**	**No**

9. Recognized International Conference(s) In Which Paper(s) Related To This Research Project Was / Were Delivered
(Please attach a copy of each conference abstract)

Month / Year / Place	**Title**	**Conference Name**	**Submitted to RGC** *(indicate the year ending of the relevant progress report)*	**Attached to this Report** *(Yes or No)*	**Acknowledged the Support of RGC** *(Yes or No)*	**Accessible from the Institutional Repository** *(Yes or No)*
12/2022/ London, UK	Optimization Methodologies for Big Data Analysis	15th International Conference of the ERCIM WG on Computational and Methodological Statistics	NA	Yes	**Yes (Annex III)**	No

10. Whether Research Experience And New Knowledge Has Been Transferred / Has Contributed To Teaching And Learning
(Please elaborate)

NA

11. Student(s) Trained
(Please attach a copy of the title page of the thesis)

Name	**Degree Registered for**	**Date of Registration**	**Date of Thesis Submission / Graduation**
NA	NA	NA	NA

12. Other Impact
(e.g. award of patents or prizes, collaboration with other research institutions, technology transfer, teaching enhancement, etc.)

NA

13. Statistics on Research Outputs

	Peer-reviewed Journal Publications	Conference Papers	Scholarly Books, Monographs and Chapters	Patents Awarded	Other Research Outputs (please specify)	
					Type	No.
No. of outputs arising directly from this research project	4 (including 1 paper under review)	1	NA	NA	NA	NA

14. Public Access Of Completion Report

(Please specify the information, if any, that cannot be provided for public access and give the reasons.)

Information that Cannot Be Provided for Public Access	Reasons
NA	NA