Multi-Block ADMM and its Convergence

Yinyu Ye
Department of Management Science and Engineering
Stanford University
Email: yyye@stanford.edu

We show that the direct extension of alternating direction method of multipliers (ADMM) with three blocks is not necessarily convergent even for solving a square system of linear equations, although its convergence proof was established 40 years ago with one or two-block. However, we prove that, in each iteration if one randomly and independently permutes the updating order of variable blocks followed by the regular multiplier update, then ADMM will converge in expectation when solving any square system of linear equations with any number of blocks. We also discuss its extension to solve general convex optimization problems, in particular, linear and quadratic programs.

Joint work with Ruoyu Sun and Tom Luo.

Optimization for High Dimensional Data

Tao Yao
Pennsylvania State University
Email: tyy1@engr.psu.edu

---

we consider folded concave penalized sparse linear regression (FCSLR), an important statistical model aimed for high-dimensional data analysis problems. Due to the presence of a folded concave penalty, the problem formulation involves both nonsmoothness and nonconvexity. We show that local solutions satisfying a second order necessary condition (SONC) entail good estimation accuracy with a probability guarantee. We discuss (1) a novel potential reduction method (PR) and (2) a new mixed integer linear programming-based global optimization (MIPGO) scheme. This talk is based on joint work with Hongcheng Liu, Runze Li and Yinyu Ye.

Efficient Approaches for Two Big Data Matrix Optimization Problems

Zhaosong Lu
Department of Mathematics
Simon Fraser University, Canada
Email: zhaosong@sfu.ca

In the first part of this talk, we consider low rank matrix completion problem, which has wide applications such as collaborative filtering, image inpainting and Microarray data imputation. In particular, we present an efficient and scalable algorithm for matrix completion. In each iteration, we pursue a rank-one matrix basis generated by the top singular vector pair of the current approximation residual and update the weights for all rank-one matrices obtained up to the current iteration. We further propose a novel weight updating rule to reduce the time and storage complexity, making the proposed algorithm scalable to large matrices. We establish a linear rate of convergence for the algorithm. Numerical experiments on many real-world large scale datasets demonstrate that our algorithm is much more efficient than the state-of-the-art algorithms while achieving similar or better prediction performance.

In the second part we consider the problem of estimating multiple graphical models simultaneously using the fused lasso penalty, which encourages adjacent graphs to share similar structures. A motivating example is the analysis of brain networks of Alzheimer's disease, which involves estimating a brain network for the normal controls (NC), a brain network for the patients with mild cognitive impairment (MCI), and a brain network for Alzheimer's patients (AD). We expect the two brain networks for NC and MCI and for MCI and AD to share common structures but not to be identical to each other. We establish a necessary and sufficient condition for the graphs to be decomposable. As a consequence, a simple but effective screening rule is proposed, which decomposes large graphs into small subgraphs and dramatically reduces the overall computational cost. Numerical experiments on both synthetic and real data demonstrate the effectiveness and efficiency of the proposed approach.

Optimal Joint Provision of Backhaul and Radio Access Networks

Zhi-quan Luo
University of Minnesota/The Chinese University of Hong Kong, Shenzhen
Email: luoxx032@umn.edu

We consider a cloud-based heterogeneous network of base stations (BSs) connected via a backhaul network of routers and wired/wireless links with limited capacity. The optimal provision of such networks requires proper resource allocation across the radio access links in conjunction with appropriate traffic engineering within the backhaul network. In this work we propose an efficient algorithm for joint resource allocation across the wireless links and the flow control over the entire network. The proposed algorithm, which maximizes the min-rate among all the transmitted commodities, is based on a decomposition approach that leverages both the Alternating Direction Method of Multipliers (ADMM) and the weighted-MMSE (WMMSE) algorithm. We show that this algorithm is easily parallelizable and converges globally to a stationary solution of the joint optimization problem. The proposed algorithm can also be extended to networks with multi-antenna nodes and other utility functions.

A Wavelet Frame Based Model for Smooth Functions and Beyond

JIANFENG CAI
Department of Mathematics
University of Iowa
Email: `jianfeng-cai@uiowa.edu`

In this talk, we present a new wavelet frame based image restoration model that explicitly treats images as piecewise smooth functions. It estimates both the image to be restored and its singularity set. It can well protect singularities, which are important image features, and provide enough regularization in smooth regions at the same time. This model penalizes the ℓ_2-norm of the wavelet frame coefficients away from the singularity set, while penalizes the ℓ_1-norm of the coefficients on the singularity set.

Some New Results on the Convergence of Multi-Block ADMM

Shiqian Ma
Department of Systems Engineering and Engineering Management
The Chinese University of Hong Kong
Email: sqma@se.cuhk.edu.hk

The alternating direction method of multipliers (ADMM) is widely used in solving structured convex optimization problems due to its superior practical performance. On the theoretical side however, a counterexample was shown by Chen et.al. indicating that the multi-block ADMM for minimizing the sum of N ($N \geq 3$) convex functions with N block variables linked by linear constraints may diverge. It is therefore of great interest to investigate further sufficient conditions on the input side which can guarantee convergence for the multi-block ADMM. In this talk, we present the following two new results regarding to the convergence of multi-block ADMM.

(1) The existing results on sufficient conditions for guaranteeing convergence of multi-block ADMM typically require the strong convexity on parts of the objective. In our work, we show convergence and convergence rate results for the multi-block ADMM applied to solve certain N-block ($N \geq 3$) convex minimization problems without requiring strong convexity. Specifically, we prove that the multi-block ADMM returns an ϵ-optimal solution within $O(1/\epsilon^2)$ iterations by solving an associated perturbation to the original problem; and the multi-block ADMM returns an ϵ-optimal solution within $O(1/\epsilon)$ iterations when it is applied to solve a certain sharing problem, under the condition that the augmented Lagrangian function satisfies the Kurdyka-Lojasiewicz property, which essentially covers most convex optimization models except for some pathological cases.

(2) It is known that the 2-block ADMM globally converges with any penalty parameter, i.e., it is a parameter free algorithm. In our work, we show that the unmodified 3-block ADMM is also parameter free, when it is applied to solving a certain sharing problem, which covers many interesting applications.

The Linearized Proximal Algorithm for Convex Composite Optimization with Applications

CHONG LI
Department of Mathematics
Zhejiang University, China
Email: cli@zju.edu.cn

We propose a linearized proximal algorithm (LPA) to solve a convex composite optimization problem. Each iteration of the LPA is a proximal minimization on the composition of the outer function and the linearization of the inner function at current iterate. The LPA has the attractive computational advantage in that the solution of each subproblem is a singleton, which avoids the difficulty of finding the whole solution set of the subproblem, as in the Gauss-Newton method (GNM), while it still maintains the same local convergence rate as that of the GNM. Under the assumptions of local weak sharp minima of order p ($p \in [1, 2]$) and the quasi-regularity condition, we establish the local superlinear convergence rate for the LPA. We also propose a globalization strategy for LPA based on the backtracking line-search and the inexact LPA, as well as the superlinear convergence results. We further apply the LPA to solve a feasibility problem, as well as a sensor network localization problem. Our numerical results illustrate that the LPA meets the demand for a robust and efficient algorithm for the sensor network localization problem.

Douglas-Rachford Splitting for Nonconvex Feasibility Problems

TING KEI PONG
Department of Applied Mathematics
The Hong Kong Polytechnic University
Email: tk.pong@polyu.edu.hk

The Douglas-Rachford (DR) splitting method is a popular approach for finding a point in the intersection of two closed convex sets and has also been applied very successfully to various nonconvex instances, even though the theoretical justification in this latter setting is far from being complete. In this talk, we aim at understanding the behavior of the DR splitting in finding an intersection point of a closed convex set C and a possibly nonconvex closed set D. In particular, we adapt this method to minimize the square distance to C subject to D. When the stepsize is small and either C or D is compact, we show that the sequence generated is bounded and any cluster point gives a stationary point of the minimization problem. Moreover, if C and D are in addition semi-algebraic, then the whole sequence is convergent. We also discuss a generalization of the method to minimize the sum of a Lipschitz differentiable function and a proper closed function, both possibly nonconvex. We present preliminary numerical results comparing the DR splitting against the alternating projection method.

This is joint work with Guoyin Li.

Newton-Krylov Methods for Problems with Embedded Monte Carlo Simulations

TIM KELLEY
Department of Mathematics
North Carolina State University, USA
Email: tim_kelley@ncsu.edu

We analyze the behavior of inexact Newton methods for problems where the nonlinear residual, Jacobian, and Jacobian-vector products are the outputs of Monte Carlo simulations. We propose algorithms which account for the randomness in the iteration, develop theory for the behavior of these algorithms, and illustrate the results with an example from neutronics.

On Game Problems with Vector-Valued Payoffs

Guangya Chen
Chinese Academy of Sciences
Email: chengy@amss.ac.cn

In this talk we introduce Shapley's model on game problems with vector-valued payoffs and Aumann's extension. We introduce also Aumann's works on utility functions and applications in research of this kind of game problems.

On Power Penalty Methods for Linear Complementarity Problems Arising from American Option Pricing

Xiaoqi Yang
Department of Applied Mathematics
The Hong Kong Polytechnic University
Email: xiao.qi.yang@polyu.edu.hk

Power penalty methods for solving a linear parabolic complementarity problem arising from American option pricing have attracted much attention. These methods require us to solve a series of systems of nonlinear equations (called penalized equations). In this paper, we first study the relationships among the solutions of penalized equations under appropriate conditions. Additionally, since these penalized equations are neither smooth nor convex, some existing algorithms, such as Newton method, cannot be applied directly to solve them. We shall apply the nonlinear Jacobian method to solve penalized equations and verify that the iteration sequence generated by the method converges monotonically to the solution of the penalized equation. Some numerical results confirm the theoretical results and the efficiency of the proposed algorithm.

On the Linear Convergence Rate of a Generalized Proximal Point Algorithm

Xiaoming Yuan
Department of Mathematics
Hong Kong Baptist University
Email: xmyuan@hkbu.edu.hk

The proximal point algorithm (PPA) has been well studied in the literature. In particular, its linear convergence rate has been studied by Rockafellar in 1976 under certain condition. We consider a generalized PPA in the generic setting of finding a zero point of a maximal monotone operator, and show that the condition proposed by Rockafellar can also sufficiently ensure the linear convergence rate for this generalized PPA. Both the exact and inexact versions of this generalized PPA are discussed. The motivation to consider this generalized PPA is that it includes as special cases the relaxed versions of some splitting methods that are originated from PPA. Thus, linear convergence results of this generalized PPA can be used to better understand the convergence of some widely usedalgorithms in the literature. We focus on the particular convex minimization context and specify Rockafellar's condition to see how to ensure the linear convergence rate for some efficient numerical schemes, including the classical augmented Lagrangian method proposed by Hensen and Powell in 1969 and its relaxed version, the original alternating direction method of multipliers (ADMM) by Glowinski and Marrocco in 1975 and its relaxed version (i.e., the generalized ADMM by Eckstein and Bertsekas in 1992). Some refined conditions weaker than existing ones are proposed in these particular contexts.

Sparse Solutions of Linear Complementarity Problems

Shuhuang Xiang
School of Mathematics and Statistics
Central South University, China
Email: xiangsh@csu.edu.cn

We consider the characterization and computation of sparse solutions and least-p-norm ($0 < p < 1$) solutions of the linear complementarity problems LCP(q, M). We show that the number of non-zero entries of any least-p-norm solution of the LCP(q, M) is less than or equal to the rank of M for any arbitrary matrix M and any number $p \in (0, 1)$, and there is $\bar{p} \in (0, 1)$ such that all least-p-norm solutions for $p \in (0, \bar{p})$ are sparse solutions. Moreover, we provide conditions on M such that a sparse solution can be found by solving convex minimization.

Applications of Sparse Solutions of Linear Complementarity Problems

Xiaojun Chen
Department of Applied Mathematics
The Hong Kong Polytechnic University
Email: xiaojun.chen@polyu.edu.hk

We consider applications of sparse solution of the LCP to the problem of portfolio selection within the Markowitz mean-variance framework. We test this methodology on the data sets constructed by Fama and French and the newly launched Shanghai-Hong Kong Stock Connect scheme. We compare with the $1/N$ strategy, and L_1 regularized portfolios in the terms of Sharpe ratio, VaR and CVaR.