2019 IEEE INTERNATIONAL WORKSHOP ON MACHINE LEARNING FOR SIGNAL PROCESSING, OCT. 13–16, 2019, PITTSBURGH, PA, USA

ON CONVERGENCE OF PROJECTED GRADIENT DESCENT FOR MINIMIZING A LARGE-SCALE QUADRATIC OVER THE UNIT SPHERE

Trung Vu, Raviv Raich, and Xiao Fu

School of EECS, Oregon State University, Corvallis, OR 97331-5501 USA

{vutru, raich, xiao.fu}@oregonstate.edu

ABSTRACT

Unit sphere-constrained quadratic optimization has been studied extensively over the past decades. While state-of-art algorithms for solving this problem often rely on relaxation or approximation techniques, there has been little research into scalable first-order methods that tackle the problem in its original form. These first-order methods are often more well-suited for the big data setting. In this paper, we provide a novel analysis of the simple projected gradient descent method for minimizing a quadratic over a sphere. When the gradient step size is sufficiently small, we show that convergence is locally linear and provide a closed-form expression for the rate. Moreover, a careful selection of the step size can stimulate convergence to the global solution while preventing convergence to local minima.

Index Terms— large-scale optimization, unit-norm constraint, quadratic programming, convergence analysis

1. INTRODUCTION

This paper studies the problem of minimizing a quadratic with a norm constraint:

$$\min_{\boldsymbol{x}\in\mathbb{R}^n} \frac{1}{2}\boldsymbol{x}^T\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}^T\boldsymbol{x} \qquad \text{subject to } \|\boldsymbol{x}\| \le 1, \qquad (1)$$

where $\boldsymbol{A} \in \mathbb{R}^{n\times n}$, $\boldsymbol{b} \in \mathbb{R}^n$ and $\|\cdot\|$ is the Euclidean norm.[1] This optimization problem arises frequently in many machine learning and signal processing applications including contour grouping [1], graph partitioning [2] and seismic inversion [3].

If the global solution $\boldsymbol{x}_\star$ of (1) lies in the interior of the unit sphere, i.e., $\|\boldsymbol{x}_\star\| < 1$, then $\boldsymbol{x}_\star$ is also the solution of the unconstrained problem. Thus, it is more challenging to consider the case when $\|\boldsymbol{x}_\star\| = 1$. To that end, we restrict our interest to the following problem of minimizing a quadratic over a sphere:

$$\min_{\boldsymbol{x}\in\mathbb{R}^n} \frac{1}{2}\boldsymbol{x}^T\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}^T\boldsymbol{x} \qquad \text{subject to } \|\boldsymbol{x}\|^2 = 1, \qquad (2)$$

In this formulation, we assume that $\boldsymbol{A}$ is symmetric, but not necessarily positive semidefinite. Hence, the objective function is potentially non-convex. Additionally, the norm constraint is non-convex. Both (1) and (2) are instances of quadratic constrained quadratic program with only one constraint (QCQP-1) and they have been extensively studied in the literature. State-of-art methods for solving this type of QCQP-1 problems in polynomial time include semidefinite relaxation (SDR) [4] and Lagrangian relaxation [5]. However, the problem size for these methods often grows quadratically, making them inapplicable to large-scale problems.

From a different standpoint, problems (1) and (2) also arise in linear algebra and optimization as the trust-region subproblem. There have been a few extensions to large-scale settings. In [6], Golub and von Matt leveraged the theory of Gauss quadrature and proposed a method to approximately solve (1) by tridiagonalizing $\boldsymbol{A}$ using the Lanczos process. In another approach, Sorensen [7] recast the trust-region subproblem in terms of a parameterized eigenvalue problem and developed an implicitly restarted Lanczos method. Related schemes can also be found in [8, 9]. In 2001, Hager introduced sequential subspace method (SSM) [10, 11], carrying out the minimization over a sequence of subspaces that are adjusted after each sequential quadratic programming (SQP) iterate. Similar to the aforementioned methods, SSM relies on Lanczos process to compute the smallest eigenvalue and corresponding eigenvector of $\boldsymbol{A}$.

In this paper, we focus on scalable first-order methods for solving (2) directly. We leverage the use of simple gradient projection and establish convergence results in the non-convex setting of the spherically constrained quadratic minimization problem, where most convergence guarantees in convex optimization start to break. Our analysis provides a

This work is partially supported by the National Science Foundation grant CCF-1254218.

[1] Generally, we can always assume $\boldsymbol{A}$ to be symmetric. Otherwise, one can define an equivalent objective function using $\hat{\boldsymbol{A}} = \frac{1}{2}(\boldsymbol{A} + \boldsymbol{A}^T)$.

978-1-7281-0824-7/19/$31.00 ©2019 IEEE

novel insight into behaviours of the algorithm in the neighborhoods of the local optima. Understanding these convergence properties enables us to (i) accommodate acceleration near the optimum—e.g., by using optimal step size selection or momentum methods—and (ii) identify ways of enabling convergence to global solution. Finally, we present numerical results that illustrate the theory developed in the paper.

2. SOLUTION PROPERTIES

Consider the Lagrange function

$$\mathcal{L}(\boldsymbol{x}, \gamma) = \frac{1}{2}\boldsymbol{x}^T\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}^T\boldsymbol{x} - \frac{1}{2}\gamma(\|\boldsymbol{x}\|^2 - 1)$$

where γ is the Lagrange multiplier. The first-order Lagrangian conditions for optimality can be specified as

$$\begin{cases} \nabla_{\boldsymbol{x}}\mathcal{L}(\boldsymbol{x}, \gamma) = \boldsymbol{A}\boldsymbol{x} - \boldsymbol{b} - \gamma\boldsymbol{x} = \boldsymbol{0}, \\ \nabla_{\gamma}\mathcal{L}(\boldsymbol{x}, \gamma) = \|\boldsymbol{x}\|^2 - 1 = 0. \end{cases}$$

For notational simplicity, we denote the residual by $\boldsymbol{r} = \boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}$ and the unit sphere by $\mathcal{S}^{n-1} = \{\boldsymbol{x} \in \mathbb{R}^n : \|\boldsymbol{x}\| = 1\}$. Formally, these conditions are given in the following lemma.

Lemma 1 (Stationary conditions)**.** *The vector $\boldsymbol{x}_*$ is a stationary point of problem (2) if and only if $\boldsymbol{x}_* \in \mathcal{S}^{n-1}$ and there exists a constant $\gamma(\boldsymbol{x}_*)$ such that $\boldsymbol{r}_* = \boldsymbol{A}\boldsymbol{x}_* - \boldsymbol{b} = \gamma(\boldsymbol{x}_*) \cdot \boldsymbol{x}_*$.*

For the rest of this manuscript, we use the shorthand notation γ to refer $\gamma(\boldsymbol{x}_*)$. Lemma 1 also implies $\gamma = \boldsymbol{r}_*^T\boldsymbol{x}_*$. Denote $\boldsymbol{P}_{\boldsymbol{x}_*}^{\perp} = \boldsymbol{I} - \boldsymbol{x}_*\boldsymbol{x}_*^T$. Let $\lambda_n = 0$ be the zero eigenvalue corresponding to the eigenvector $\boldsymbol{x}_*$ of the matrix $\boldsymbol{P}_{\boldsymbol{x}_*}^{\perp}\boldsymbol{A}\boldsymbol{P}_{\boldsymbol{x}_*}^{\perp}$ and $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_{n-1}$ be the remaining $n-1$ eigenvalues. Noticeably, the eigenvalues of $\boldsymbol{P}_{\boldsymbol{x}_*}^{\perp}\boldsymbol{A}\boldsymbol{P}_{\boldsymbol{x}_*}^{\perp}$ can be bounded by

$$\lambda_{\min}(\boldsymbol{A}) \leq \lambda_{n-1} \leq \ldots \leq \lambda_1 \leq \lambda_{\max}(\boldsymbol{A}),$$

where $\lambda_{\min}(\boldsymbol{A})$ and $\lambda_{\max}(\boldsymbol{A})$ are the smallest and largest eigenvalues of $\boldsymbol{A}$, respectively. Moreover, the relationship among those eigenvalues and the Lagrange multiplier provides necessary and sufficient conditions for determining the type of a stationary point.

Lemma 2. *A stationary point $\boldsymbol{x}_*$ of problem (2) is a strict local minimum if and only if $\gamma(\boldsymbol{x}_*) < \lambda_{n-1}(\boldsymbol{x}_*)$. Furthermore, $\boldsymbol{x}_\star$ is a **global minimizer** of problem (2) if and only if $\gamma(\boldsymbol{x}_\star) \leq \lambda_{\min}(\boldsymbol{A})$.*

Due to space limitation, we provide proofs of our lemmas at https://trungvietvu.github.io/files/MLSP19_Appendix.pdf.

Example 1. *Figure 1 demonstrates various cases where there are different numbers of stationary points. As an exemplification, let us examine the derivation of the problem in Fig. 1(b):*

$$\min_{x_1, x_2} \frac{1}{2}(4x_1^2 + x_2^2) - 2x_1 \quad \textit{s.t. } x_1^2 + x_2^2 = 1.$$

For each stationary point $\boldsymbol{x}_*$, the matrix $\boldsymbol{P}_{\boldsymbol{x}_*}^{\perp}\boldsymbol{A}\boldsymbol{P}_{\boldsymbol{x}_*}^{\perp}$ has one zero eigenvalue corresponding to the eigenvector $\boldsymbol{x}_*$, and the other non-zero eigenvalue $\lambda_1 = \lambda_{n-1}$ (since $n = 2$) lies between $\lambda_{\min}(\boldsymbol{A}) = 1$ and $\lambda_{\max}(\boldsymbol{A}) = 4$. Omitting the detailed calculation, we list the four stationary points of this problem as follows: (i) a global maximum at $[x_1, x_2] = [-1, 0]$ with $\gamma = 6, \lambda_1 = 1$; (ii) a local maximum at $[x_1, x_2] = [1, 0]$ with $\gamma = 2, \lambda_1 = 1$; and (iii) 2 local (also global) minima at $[x_1, x_2] = [\frac{2}{3}, \pm\frac{\sqrt{5}}{3}]$ with $\gamma = 1, \lambda_1 = \frac{8}{3}$.

3. THE PROJECTED GRADIENT ALGORITHM

The projected gradient descent approach (see Algorithm 1) starts at an initial point $x^{(0)}$, then performs the update

$$\boldsymbol{x}^{(t+1)} = \boldsymbol{f}_\alpha(\boldsymbol{x}^{(t)}) = \mathcal{P}_{\mathcal{S}^{n-1}}\Big(\boldsymbol{x}^{(t)} - \alpha(\boldsymbol{A}\boldsymbol{x}^{(t)} - \boldsymbol{b})\Big), \quad (3)$$

where $\alpha > 0$ is the step size and $\mathcal{P}_{\mathcal{S}^{n-1}}(\cdot) : \mathbb{R}^n \to \mathbb{R}^n$ is the spherical projection uniquely given by

$$\mathcal{P}_{\mathcal{S}^{n-1}}(\boldsymbol{x}) = \begin{cases} \frac{\boldsymbol{x}}{\|\boldsymbol{x}\|} & \text{if } \boldsymbol{x} \neq \boldsymbol{0}, \\ \boldsymbol{e} & \text{if } \boldsymbol{x} = \boldsymbol{0}, \end{cases}$$

with $\boldsymbol{e} \in \mathcal{S}^{n-1}$ such that $\boldsymbol{e}$ and $\boldsymbol{A}\boldsymbol{e} - \boldsymbol{b}$ are not collinear. The definition of projection at $\boldsymbol{0}$ is just for numerical issues when the algorithm encounters the origin at some iteration. In practice, we can choose $\boldsymbol{e}$ to be one of the natural basis, i.e., $[1, 0, \ldots, 0]$. Next, let us consider some important properties associated with Algorithm 1.

Definition 1. *A fixed point of $\boldsymbol{f}_\alpha$ is defined as any vector $\bar{\boldsymbol{x}} \in \mathbb{R}^n$ such that*

$$\boldsymbol{f}_\alpha(\bar{\boldsymbol{x}}) = \mathcal{P}_{\mathcal{S}^{n-1}}\big(\bar{\boldsymbol{x}} - \alpha(\boldsymbol{A}\bar{\boldsymbol{x}} - \boldsymbol{b})\big) = \bar{\boldsymbol{x}}.$$

Lemma 3. *The vector $\bar{\boldsymbol{x}}$ is a fixed point of $\boldsymbol{f}_\alpha$ if and only if $\bar{\boldsymbol{x}} \in \mathcal{S}^{n-1}$ and there exists a constant $\gamma < \frac{1}{\alpha}$ such that $\bar{\boldsymbol{r}} = \boldsymbol{A}\bar{\boldsymbol{x}} - \boldsymbol{b} = \gamma\bar{\boldsymbol{x}}$.*

Proof. Since $\bar{\boldsymbol{x}}$ is a fixed point of $\boldsymbol{f}_\alpha$, we have

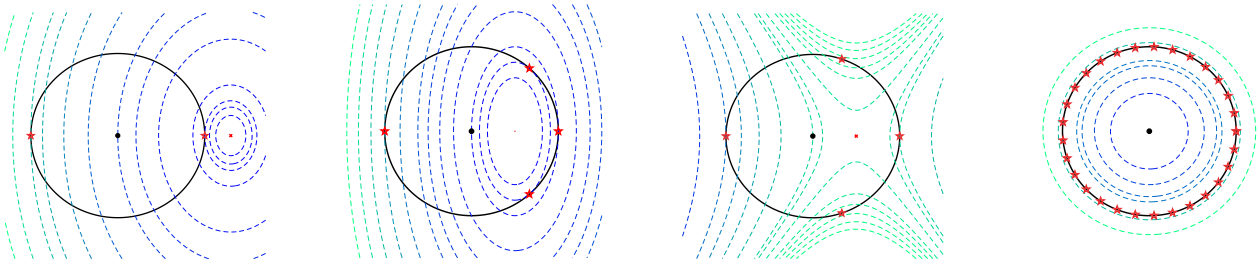
$$\bar{\boldsymbol{x}} = \boldsymbol{f}_\alpha(\bar{\boldsymbol{x}}) = \mathcal{P}_{\mathcal{S}^{n-1}}\big(\bar{\boldsymbol{x}} - \alpha(\boldsymbol{A}\bar{\boldsymbol{x}} - \boldsymbol{b})\big).$$

Consequently, $\bar{\boldsymbol{x}} \in \mathcal{S}^{n-1}$. Furthermore, if $\bar{\boldsymbol{x}} - \alpha(\boldsymbol{A}\bar{\boldsymbol{x}} - \boldsymbol{b}) = \boldsymbol{0}$, then $\bar{\boldsymbol{x}} = \mathcal{P}_{\mathcal{S}^{n-1}}(\boldsymbol{0}) = \boldsymbol{e}$. But this contradicts with the non-collinearity of $\boldsymbol{e}$ and $\boldsymbol{A}\boldsymbol{e} - \boldsymbol{b}$. Thus, it must be the case that $\bar{\boldsymbol{x}} - \alpha(\boldsymbol{A}\bar{\boldsymbol{x}} - \boldsymbol{b}) \neq \boldsymbol{0}$, and hence $\bar{\boldsymbol{x}} = \frac{\bar{\boldsymbol{x}} - \alpha\bar{\boldsymbol{r}}}{\|\bar{\boldsymbol{x}} - \alpha\bar{\boldsymbol{r}}\|}$. There exists a constant γ such that $\bar{\boldsymbol{r}} = \gamma\bar{\boldsymbol{x}}$. Substituting back into the fixed-point equation yields

$$\bar{\boldsymbol{x}} = \frac{(1 - \alpha\gamma)\bar{\boldsymbol{x}}}{|1 - \alpha\gamma|\,\|\bar{\boldsymbol{x}}\|} = \frac{1 - \alpha\gamma}{|1 - \alpha\gamma|}\frac{\bar{\boldsymbol{x}}}{\|\bar{\boldsymbol{x}}\|} = \text{sign}(1 - \alpha\gamma) \cdot \bar{\boldsymbol{x}}.$$

Therefore, the sign of $1 - \alpha\gamma$ must be 1 or $1 - \alpha\gamma > 0$. □

From Lemma 2 and 3, we can establish the necessary and sufficient conditions for a fixed point of $\boldsymbol{f}_\alpha$ to be a stationary point as follows.



(a) $\boldsymbol{A} = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}, \boldsymbol{b} = \begin{bmatrix} 3 \\ 0 \end{bmatrix}$ (b) $\boldsymbol{A} = \begin{bmatrix} 4 & 0 \\ 0 & 1 \end{bmatrix}, \boldsymbol{b} = \begin{bmatrix} 2 \\ 0 \end{bmatrix}$ (c) $\boldsymbol{A} = \begin{bmatrix} -2 & 0 \\ 0 & 1 \end{bmatrix}, \boldsymbol{b} = \begin{bmatrix} -1 \\ 0 \end{bmatrix}$ (d) $\boldsymbol{A} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \boldsymbol{b} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

Fig. 1: Examples of minimizing a quadratic over a sphere. Stationary points are given in red stars. In 2D scenario, they can be either local minima or local maxima (a-c). It is also possible that a local optimum lies in a continuum of optima (d).

Algorithm 1: Projected Gradient Descent (PGD)

1: Initialize $\boldsymbol{x}^{(0)}$
2: **for** $t = 0, 1, \ldots$ **do**
3: $\quad \boldsymbol{z}^{(t+1)} = \boldsymbol{x}^{(t)} - \alpha(\boldsymbol{A}\boldsymbol{x}^{(t)} - \boldsymbol{b})$
4: $\quad \boldsymbol{x}^{(t+1)} = \mathcal{P}_{\mathcal{S}^{n-1}}(\boldsymbol{z}^{(t+1)})$

Corollary 1. *The vector $\boldsymbol{x}_*$ is a stationary point of problem (2) if and only if there exists $\alpha > 0$ such that $\boldsymbol{x}_*$ is a fixed point of $\boldsymbol{f}_\alpha$.*

Example 2. *Continued from Example 1, we illustrate fixed points with different step sizes in Fig. 2. When α is small enough, all stationary points can be fixed points. As α increases, only stationary points with the multiplier $\gamma < 1/\alpha$ remains to be fixed points of $\boldsymbol{f}_\alpha$. Interestingly, while any convergence point of Algorithm 1 with step size α is a fixed point of the iterated function $\boldsymbol{f}_\alpha$, the vice versa is not true: the global maximum at $[x_1, x_2] = [-1, 0]$ is a fixed point of $\boldsymbol{f}_\alpha$, for $\alpha < 1/6$, but as can be seen in the next section, it is not a convergence point of the algorithm.*

4. CONVERGENCE ANALYSIS

In this section, we present our result on the local uniform convergence of Algorithm 1 with a certain choice of step size to a strict local optimum. The convergence is shown to be linear and the asymptotic rate is given in a closed-form expression. The challenges come the non-convexity of the norm constraint and (potentially) the negative curvature of the objective function. Let us begin with the analysis of the projection operator.

Lemma 4 (Taylor series expansion of the projection)**.** *Let $\boldsymbol{x} \in \mathbb{R}^n$ be a nonzero vector and $\boldsymbol{\delta}$ be a small perturbation such that $\|\boldsymbol{\delta}\| \ll \|\boldsymbol{x}\|$. Then,*

$$\mathcal{P}_{\mathcal{S}^{n-1}}(\boldsymbol{x} + \boldsymbol{\delta}) = \mathcal{P}_{\mathcal{S}^{n-1}}(\boldsymbol{x}) + \tfrac{1}{\|\boldsymbol{x}\|}\Big(\boldsymbol{I} - \tfrac{\boldsymbol{x}\boldsymbol{x}^T}{\|\boldsymbol{x}\|^2}\Big)\boldsymbol{\delta} + O\big(\|\boldsymbol{\delta}\|^2\big).$$

Now, considering the convergence of Algorithm 1 in the region near a strict local minimum $\boldsymbol{x}_*$ where $\boldsymbol{r}_* = \boldsymbol{A}\boldsymbol{x}_* - \boldsymbol{b} = \gamma\boldsymbol{x}_*$. Denote $\boldsymbol{\delta}^{(t)} = \boldsymbol{x}^{(t)} - \boldsymbol{x}_*$ and reorganize the update equation (3) as

$$\begin{aligned}\boldsymbol{\delta}^{(t+1)} &= \mathcal{P}_{\mathcal{S}^{n-1}}\big(\boldsymbol{x}_* - \alpha(\boldsymbol{A}\boldsymbol{x}_* - \boldsymbol{b}) + (\boldsymbol{I} - \alpha\boldsymbol{A})\boldsymbol{\delta}^{(t)}\big) - \boldsymbol{x}_* \\ &= \mathcal{P}_{\mathcal{S}^{n-1}}\big((1-\alpha\gamma)\boldsymbol{x}_* + (\boldsymbol{I} - \alpha\boldsymbol{A})\boldsymbol{\delta}^{(t)}\big) - \mathcal{P}_{\mathcal{S}^{n-1}}\big((1-\alpha\gamma)\boldsymbol{x}_*\big).\end{aligned}$$

Substituting $\boldsymbol{x} = (1-\alpha\gamma)\boldsymbol{x}_*$ and $\boldsymbol{\delta} = (\boldsymbol{I} - \alpha\boldsymbol{A})\boldsymbol{\delta}^{(t)}$ into Lemma 4 yields

$$\begin{aligned}\boldsymbol{\delta}^{(t+1)} = \frac{1}{\|(1-\alpha\gamma)\boldsymbol{x}_*\|}\Big(\boldsymbol{I} - \frac{(1-\alpha\gamma)\boldsymbol{x}_*(1-\alpha\gamma)\boldsymbol{x}_*^T}{\|(1-\alpha\gamma)\boldsymbol{x}_*\|^2}\Big) \\ \cdot (\boldsymbol{I} - \alpha\boldsymbol{A})\boldsymbol{\delta}^{(t)} + O\big(\|\boldsymbol{\delta}^{(t)}\|^2\big).\end{aligned}$$

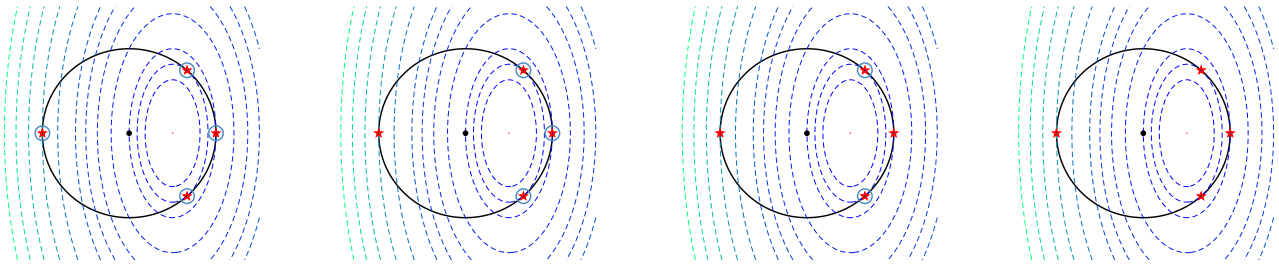
Assume the step size is chosen such that $\alpha\gamma < 1$, and recall that $\boldsymbol{P}_{\boldsymbol{x}_*}^{\perp} = \boldsymbol{I} - \boldsymbol{x}_*\boldsymbol{x}_*^T$ for $\|\boldsymbol{x}_*\| = 1$. The recursion can be rewritten as

$$\boldsymbol{\delta}^{(t+1)} = \frac{1}{1-\alpha\gamma}\boldsymbol{P}_{\boldsymbol{x}_*}^{\perp}(\boldsymbol{I} - \alpha\boldsymbol{A})\boldsymbol{\delta}^{(t)} + O\big(\|\boldsymbol{\delta}^{(t)}\|^2\big). \quad (4)$$

The stability of a general nonlinear difference equation of the form $x^{k+1} = Tx^k + o(\|x^k\|)$ has been well-studied in [12, 13]. In particular, let ρ_α be the spectral radius of $(1-\alpha\gamma)^{-1}\boldsymbol{P}_{\boldsymbol{x}_*}^{\perp}(\boldsymbol{I} - \alpha\boldsymbol{A})$, i.e., the largest absolute value of its eigenvalues. If $\rho_\alpha < 1$, then the series $\{\boldsymbol{\delta}^{(t)}\}$ approaches zeros with sufficiently small $\boldsymbol{\delta}^{(0)}$, where

$$\|\boldsymbol{\delta}^{(t)}\| \le K\|\boldsymbol{\delta}^{(0)}\|\big(\rho_\alpha + o(1)\big)^t.$$

Therefore, to prove the local convergence of the PGD algorithm, it is sufficient to show that $\rho_\alpha < 1$. We present our main result on the local uniform convergence to a strict local minimum as follows.



(a) $0 < \alpha < 1/6$ (b) $1/6 < \alpha < 1/2$ (c) $1/2 < \alpha < 1$ (d) $\alpha > 1$

Fig. 2: Stationary points (red stars) versus fixed points (blue circles) with different step size α in optimizing the quadratic objective $\frac{1}{2}(4x_1^2 + x_2^2) - 2x_1$ over the unit circle. Dashed lines are the contour levels of the objective value.

Definition 2. *Algorithm 1 with step size α converges **locally uniformly** to $\boldsymbol{x}_*$ if and only if there exists a constant ϵ such that for any $\boldsymbol{x}^{(0)}$ satisfying $\left\|\boldsymbol{x}^{(0)} - \boldsymbol{x}_*\right\| \leq \epsilon$, we have $\left\|\boldsymbol{x}^{(t)} - \boldsymbol{x}_*\right\| \leq \epsilon$, $\forall t = 0, 1, \ldots$ and $\lim_{t\to\infty} \left\|\boldsymbol{x}^{(t)} - \boldsymbol{x}_*\right\| = 0$.*

Theorem 1. *The vector $\boldsymbol{x}_*$ is a strict local minimum of problem (2), i.e. $\gamma < \lambda_{n-1}$, if and only if there exists $\alpha > 0$ such that Algorithm 1 with step size α converges locally uniformly to $\boldsymbol{x}_*$. Furthermore, for any step size $\alpha > 0$ such that $\alpha(\lambda_1 + \gamma) < 2$, the sequence $\{\boldsymbol{x}^{(t)}\}$ satisfies*

$$\|\boldsymbol{x}^{(t)} - \boldsymbol{x}_*\| \leq K\|\boldsymbol{x}^{(0)} - \boldsymbol{x}_*\|\big(\rho_\alpha + o(1)\big)^t,$$

for some constant $K > 0$ and $\rho_\alpha = \max_{1\leq i\leq n-1} \frac{|1-\alpha\lambda_i|}{1-\alpha\gamma}$.

The proof of Theorem 1 is given in the appendix. The theorem reveals PGD converges to a local minimum at an asymptotic linear rate ρ_α. Note that in our problem, $\boldsymbol{A}$ is not necessarily PSD, meaning λ_i could be negative. To facilitate acceleration, one can speed up the convergence by optimizing over the step size α.

Lemma 5. *The optimal rate of local convergence and the optimal step size for Algorithm 1 are given by*

$$\begin{cases} \rho_* = \frac{\lambda_1 - \lambda_{n-1}}{\lambda_1 + \lambda_{n-1} - 2\gamma}, & \alpha_* = \frac{2}{\lambda_1 + \lambda_{n-1}} \quad \text{if } \lambda_1 + \lambda_{n-1} > 0 \\ \rho_* = \frac{\lambda_{n-1}}{\gamma}, & \alpha_* = \infty \quad \text{otherwise.} \end{cases}$$

Example 3. *Continued from Example 2, Theorem 1 states that the PGD algorithm only converges locally uniformly to the two local minima at $[2/3, \pm\sqrt{(5)}/3]$ with $\lambda_1 = \frac{8}{3}$. Notice that these points are fixed points of $\boldsymbol{f}_\alpha$ for $\alpha < 1$. Since $\lambda_1 = \lambda_{n-1}$, the optimal rate is $\rho_* = 0$ with step size $\alpha_* = \frac{3}{8}$. In this case, the convergence is quadratic due to the residual term in (4).*

Global convergence of the PGD algorithm. Theorem 1 also implies that for any step size α satisfying $\rho_\alpha > 1$, the algorithm tends to move away from the local minimum $\boldsymbol{x}_*$. This intuition leads us to the following strategy for step size selection: choosing α large enough such that $g(\alpha, \boldsymbol{x}_*) \geq 2$, where $g(\alpha, \boldsymbol{x}_*) \triangleq \alpha(\lambda_1(\boldsymbol{x}_*) + \gamma(\boldsymbol{x}_*))$, for all strict local minimum $\boldsymbol{x}_*$ except the global minimum $\boldsymbol{x}_\star$.

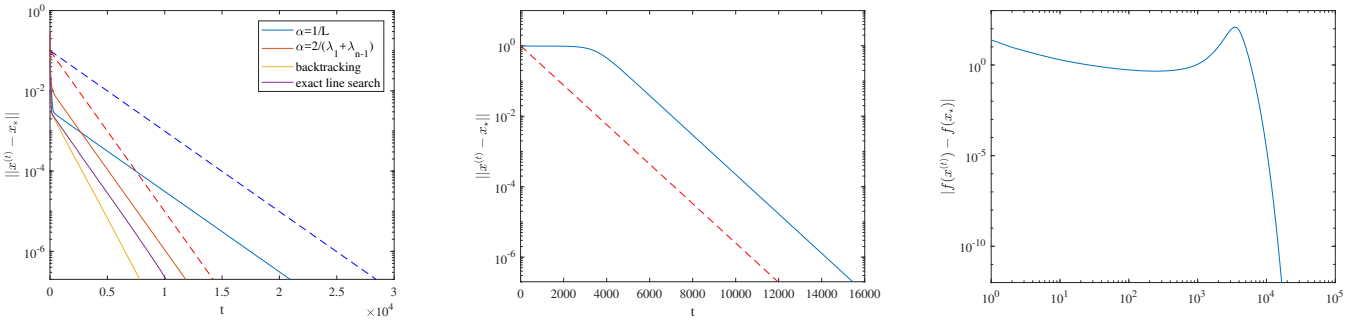
Remark 1. *Assume that there exists sufficiently large α satisfying $g(\alpha, \boldsymbol{x}_\star) < 2$ for any global minimum $\boldsymbol{x}_\star$ and $g(\alpha, \boldsymbol{x}_*) \geq 2$ for any strict local minimum $\boldsymbol{x}_*$. Then Algorithm 1 with step size α converges to one of the optimal solutions $\boldsymbol{x}_\star$ at an asymptotic geometric rate of $\rho_\alpha(\boldsymbol{x}_\star)$.*

5. NUMERICAL RESULTS

Motivating applications of problem (1) are the well-known trust-region subproblem in nonlinear optimization [14], the variational problem in structural limit analysis [15], and the optimizing precoder method in transmitter based CDMA optimization [16]. For the purpose of demonstrating the theoretical analysis, we will focus on numerical results for local convergence of Algorithm 1 with different step sizes, and empirical evidence for our conjecture about the global convergence with an appropriate step size. In our experiment, we first generate a random symmetric matrix $\boldsymbol{A}$ of size $n = 1000$ such that the smallest eigenvalue is far away from the other eigenvalues. Then, we choose one multiplier for the global solution $\gamma(\boldsymbol{x}_\star) < \lambda_{\min}(\boldsymbol{A})$ and one multiplier for the local solution $\gamma(\boldsymbol{x}_* > \lambda_{\min}(\boldsymbol{A})$. Next, the coefficient vector $\boldsymbol{b}$ is chosen such that $\boldsymbol{b}^T(\boldsymbol{A} - \gamma(\boldsymbol{x}_\star)\boldsymbol{I})^{-2}\boldsymbol{b} = \boldsymbol{b}^T(\boldsymbol{A} - \gamma(\boldsymbol{x}_*)\boldsymbol{I})^{-2}\boldsymbol{b} = 1$. Finally, we compute $\boldsymbol{x}_\star = (\boldsymbol{A} - \gamma(\boldsymbol{x}_\star)\boldsymbol{I})^{-1}\boldsymbol{b}$ and $\boldsymbol{x}_* = (\boldsymbol{A} - \gamma(\boldsymbol{x}_*)\boldsymbol{I})^{-1}\boldsymbol{b}$.

For local convergence, starting at an initial point $\boldsymbol{x}^{(0)}$ close to the local minimum $\boldsymbol{x}_*$, we examine four PGD algorithms with different step sizes:

(i) Commonly used step size: $\alpha = 1/L$, where L is the spectral radius of $\boldsymbol{A}$. This step size selection is often used in many classic proofs of convergence in convex optimization, where an L-smooth objective function can be guaranteed to monotonically decrease through PGD iterations.



(a) (b) (c)

Fig. 3: (a) Local convergence of the projected gradient methods with different step sizes for solving a unit-constrained least squares. (b-c) Empirical evidence of the convergence to global optimum where $\boldsymbol{x}^{(0)}$ is initialized near a local optimum and α is chosen according to Remark 1. Dashed lines are added as an illustration for the theoretical bounds for the convergence rate of fixed step size methods (up to a constant).

(ii) Optimal step size: $\alpha = \alpha_*$ as in Lemma 5. We choose $\boldsymbol{A}$ and $\boldsymbol{x}_*$ such that $\lambda_1 + \lambda_{n-1} > 0$, hence $\alpha_* = 2/(\lambda_1 + \lambda_{n-1})$.
(iii) Projected backtracking line search: rewriting the PGD update as generalized gradient step $\boldsymbol{x}^{(t+1)} = \boldsymbol{x}^{(t)} - \alpha_t \boldsymbol{G}_{\alpha_t}(\boldsymbol{x}^{(t)})$, where $\boldsymbol{G}_\alpha(\boldsymbol{x}) = \frac{1}{\alpha}\Big(\boldsymbol{x} - \mathcal{P}_{\mathcal{S}^{n-1}}\big(\boldsymbol{x} - \alpha(\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b})\big)\Big)$. Denote the quadratic objective by $q(\boldsymbol{x})$. Starting with $\alpha = 1$, we shrink $\alpha = \beta\alpha$, for $0 < \beta < 1$, while

$$q\big(\boldsymbol{x} - \alpha \boldsymbol{G}_\alpha(\boldsymbol{x})\big) > q(\boldsymbol{x}) - \alpha(\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b})^T \boldsymbol{G}_\alpha(\boldsymbol{x}) + \frac{\alpha}{2}\left\|\boldsymbol{G}_\alpha(\boldsymbol{x})\right\|^2.$$

In our case, this backtracking condition can be simplified to

$$\boldsymbol{G}_\alpha(\boldsymbol{x})^T \boldsymbol{A} \boldsymbol{G}_\alpha(\boldsymbol{x}) > \alpha \left\|\boldsymbol{G}_\alpha(\boldsymbol{x})\right\|^2.$$

(iv) Exact line search: finding the step size that maximizes the decrease in objective function

$$\alpha_{\min} = \min_{\alpha > 0} q\big(\mathcal{P}_{\mathcal{S}^{n-1}}(\boldsymbol{x} - \alpha(\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}))\big).$$

As can be seen from Fig. 3(a), the convergence of PGD with step size $\alpha = 1/L$ (blue) is the slowest among the considered methods, followed by the one with optimal step size $\alpha = 2/(\lambda_1 + \lambda_{n-1})$ (red). Note that they both match the asymptotic rate predicted in theory. The adaptive schemes, namely projected backtracking (yellow) and exact line search (magenta) perform slightly better than the optimal fixed step size scheme.

For global convergence, we purposely initialize the algorithm at the same $\boldsymbol{x}^{(0)}$ that is close to the **local** minimum $\boldsymbol{x}_*$, and run the PGD algorithm with step size $\alpha = \frac{1}{2}\big(\frac{2}{\lambda_1(\boldsymbol{x}_\star) + \gamma(\boldsymbol{x}_\star)} + \frac{2}{\lambda_1(\boldsymbol{x}_*) + \gamma(\boldsymbol{x}_*)}\big)$. It is easy to verify that α satisfies the condition in Remark 1: $g(\alpha, \boldsymbol{x}_*) < 2 < g(\alpha, \boldsymbol{x}_\star)$. Figure 3 demonstrates the convergence of the algorithm to the global minimizer $\boldsymbol{x}_\star$, in terms of the distance to the solution (b) and the decrease in the objective value (c). In the first 5000 iterations, the algorithm tries to escape from the local minimum. Then it experiences a period of fluctuation before actually getting attracted by the global minimum. Notice that when it reaches the neighborhood of $\boldsymbol{x}_\star$, the (monotonic) linear convergence appears. [2]

6. CONCLUSION AND FUTURE WORKS

We analyze the projected gradient descent approach to minimizing a quadratic over a sphere. We showed that the algorithm always converges linearly to a strict local minimum in its neighborhood. Further, we provide the closed-form expression for convergence rate and identify ways of achieving optimal rate of convergence near the optimum. Our analysis can be extended in the following directions: (i) minimizing a quadratic over an ellipsoid; (ii) acceleration of gradient projection using momentum; and (iii) analysis of convergence to a continuum of optima.

7. APPENDIX

The proof of Theorem 1 is given as follows.
[⇒] First, if Algorithm 1 converges locally uniformly to $\boldsymbol{x}_*$, our goal is to prove $\gamma < \lambda_{n-1}$. By contradiction, assume that $\gamma > \lambda_{n-1}$ [3]. Then choosing $\boldsymbol{x}^{(0)} = \boldsymbol{x}_* + \epsilon \boldsymbol{u}_{n-1}$, where $\boldsymbol{u}_{n-1}$ is the eigenvector corresponding to λ_{n-1}, leads to

$$\boldsymbol{\delta}^{(1)} = \frac{|1 - \alpha\lambda_{n-1}|}{1 - \alpha\gamma}\boldsymbol{\delta}^{(0)} = \frac{1 - \alpha\lambda_{n-1}}{1 - \alpha\gamma}\boldsymbol{\delta}^{(0)}$$

$$\Rightarrow \left\|\boldsymbol{x}^{(1)} - \boldsymbol{x}_*\right\| = \left\|\boldsymbol{\delta}^{(1)}\right\| = \left|\frac{1 - \alpha\lambda_{n-1}}{1 - \alpha\gamma}\right| \|\epsilon \boldsymbol{u}_{n-1}\|$$

$$= \frac{1 - \alpha\lambda_{n-1}}{1 - \alpha\gamma}\epsilon > \epsilon. \qquad (\text{since } \alpha\lambda_{n-1} < \alpha\gamma < 1)$$

[2] This result is provided merely as an illustration of a typical run, not to be considered as an empirical proof. In our experiments, we re-ran simulations multiple times with various problem sizes and always observed convergence.

[3] The case $\gamma = \lambda_{n-1}$ leads to convergence to a continuum which we will left as an open question for future development of this paper.

This contradicts with the assumption that the sequence $\{\boldsymbol{x}^{(t)}\}$ lies inside the ϵ-vicinity of $\boldsymbol{x}_*$.

[$\Leftarrow$] Conversely, we will show that if $\boldsymbol{x}_*$ is a strict local minimum, then for any $\alpha > 0$ such that $\alpha(\lambda_1 + \gamma) < 2$, Algorithm 1 with step size α converges locally uniformly to $\boldsymbol{x}_*$. By the same argument in [13], to prove the local stability of equation (4), it is sufficient to consider the linear equation without the quadratic residual

$$\boldsymbol{\delta}^{(t+1)} = \frac{1}{1-\alpha\gamma}\boldsymbol{P}_{\boldsymbol{x}_*}^{\perp}(\boldsymbol{I} - \alpha\boldsymbol{A})\delta^{(t)}.$$

The above equation implies $\boldsymbol{\delta}^{(t)} = \boldsymbol{P}_{\boldsymbol{x}_*}^{\perp}\boldsymbol{\delta}^{(t)}$ for $t = 1, 2, \ldots$. Thus, we have

$$\begin{aligned}\boldsymbol{\delta}^{(t+1)} &= \frac{\boldsymbol{P}_{\boldsymbol{x}_*}^{\perp}(\boldsymbol{I} - \alpha\boldsymbol{A})\boldsymbol{P}_{\boldsymbol{x}_*}^{\perp}}{1-\alpha\gamma}\delta^{(t)} = \frac{\boldsymbol{P}_{\boldsymbol{x}_*}^{\perp} - \alpha\boldsymbol{P}_{\boldsymbol{x}_*}^{\perp}\boldsymbol{A}\boldsymbol{P}_{\boldsymbol{x}_*}^{\perp}}{1-\alpha\gamma}\boldsymbol{\delta}^{(t)} \\ &= \frac{(\boldsymbol{I} - \alpha\boldsymbol{P}_{\boldsymbol{x}_*}^{\perp}\boldsymbol{A}\boldsymbol{P}_{\boldsymbol{x}_*}^{\perp})\boldsymbol{P}_{\boldsymbol{x}_*}^{\perp}}{1-\alpha\gamma}\boldsymbol{\delta}^{(t)} = \frac{\boldsymbol{I} - \alpha\boldsymbol{P}_{\boldsymbol{x}_*}^{\perp}\boldsymbol{A}\boldsymbol{P}_{\boldsymbol{x}_*}^{\perp}}{1-\alpha\gamma}\boldsymbol{\delta}^{(t)}.\end{aligned}$$

Now consider the matrix $\boldsymbol{P}_{\boldsymbol{x}_*}^{\perp}\boldsymbol{A}\boldsymbol{P}_{\boldsymbol{x}_*}^{\perp}$. There exists an eigenvalue decomposition $\boldsymbol{P}_{\boldsymbol{x}_*}^{\perp}\boldsymbol{A}\boldsymbol{P}_{\boldsymbol{x}_*}^{\perp} = \boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{U}^T$ where $\boldsymbol{U} = [\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_{n-1}, \boldsymbol{x}_*]$ is an orthogonal matrix and $\boldsymbol{\Lambda} = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_{n-1}, 0)$. Let $\boldsymbol{y}^{(t)} = \boldsymbol{U}^T\boldsymbol{\delta}^{(t)}$. Then

$$\boldsymbol{y}^{(t+1)} = \boldsymbol{U}^T\boldsymbol{\delta}^{(t+1)} = \tfrac{\boldsymbol{I}-\alpha\boldsymbol{\Lambda}}{1-\alpha\gamma}\boldsymbol{y}^{(t)} = \left(\tfrac{\boldsymbol{I}-\alpha\boldsymbol{\Lambda}}{1-\alpha\gamma}\right)^t\boldsymbol{y}^{(1)}. \quad (5)$$

In addition, since the last column of $\boldsymbol{U}$ is $\boldsymbol{x}_*$, we can compute the last element of $\boldsymbol{y}^{(1)}$ by

$$y_n^{(1)} = \boldsymbol{x}_*^T\boldsymbol{\delta}^{(1)} = \boldsymbol{x}_*^T(\boldsymbol{I} - \boldsymbol{x}_*\boldsymbol{x}_*^T)\boldsymbol{\delta}^{(0)} = 0. \quad (6)$$

From (5) and (6), we obtain

$$\left\|\boldsymbol{y}^{(t+1)}\right\| \leq \max_{1\leq i\leq n-1}\left|\frac{1-\alpha\lambda_i}{1-\alpha\gamma}\right|^t \cdot \left\|\boldsymbol{y}^{(1)}\right\|.$$

Since $\boldsymbol{x}_*$ is a strict local minimum, it follows from Lemma 2 that $\gamma < \lambda_{n-1} \leq \lambda_1$. Combining with the condition $\alpha\lambda_1 + \alpha\gamma < 2$, we obtain $\alpha\gamma < 1$. In order to show convergence, it remains to prove the inequality

$$\max_{1\leq i\leq n-1}\frac{|1-\alpha\lambda_i|}{1-\alpha\gamma} < 1 \Leftrightarrow |1-\alpha\lambda_i| < 1-\alpha\gamma,\ \forall\, 1 \leq i \leq n-1.$$

Indeed, this inequality stems from the fact that $\alpha\lambda_1 + \alpha\gamma < 2$ and $\gamma < \lambda_{n-1} \leq \ldots \leq \lambda_1$.

8. REFERENCES

[1] Jean Gallier, *Quadratic Optimization and Contour Grouping*, pp. 439–457, Springer New York, New York, NY, 2011.

[2] William W Hager and Yaroslav Krylyuk, "Graph partitioning and continuous quadratic programming," *SIAM Journal on Discrete Mathematics*, vol. 12, no. 4, pp. 500–523, 1999.

[3] William Menke, *Geophysical data analysis: Discrete inverse theory*, Academic press, 2018.

[4] Zhi-Quan Luo, Nicholas D Sidiropoulos, Paul Tseng, and Shuzhong Zhang, "Approximation bounds for quadratic optimization with homogeneous quadratic constraints," *SIAM Journal on optimization*, vol. 18, no. 1, pp. 1–28, 2007.

[5] Jaehyun Park and Stephen Boyd, "General heuristics for nonconvex quadratically constrained quadratic programming," *arXiv preprint arXiv:1703.07870*, 2017.

[6] Gene H Golub and Urs Von Matt, "Quadratically constrained least squares and quadratic problems," *Numerische Mathematik*, vol. 59, no. 1, pp. 561–580, 1991.

[7] Danny C Sorensen, "Minimization of a large-scale quadratic function subject to a spherical constraint," *SIAM Journal on Optimization*, vol. 7, no. 1, pp. 141–161, 1997.

[8] Franz Rendl and Henry Wolkowicz, "A semidefinite framework for trust region subproblems with applications to large scale minimization," *Mathematical Programming*, vol. 77, no. 1, pp. 273–299, 1997.

[9] Marielba Rojas, *A large-scale trust-region approach to the regularization of discrete ill-posed problems*, Ph.D. thesis, Rice University, 1999.

[10] William W Hager, "Minimizing a quadratic over a sphere," *SIAM Journal on Optimization*, vol. 12, no. 1, pp. 188–208, 2001.

[11] William Hager and Soonchul Park, "Global convergence of SSM for minimizing a quadratic over a sphere," *Mathematics of Computation*, vol. 74, no. 251, pp. 1413–1423, 2005.

[12] Richard Bellman, *Stability theory of differential equations*, Courier Corporation, 2008.

[13] B. Polyak, "Some methods of speeding up the convergence of iteration methods," *USSR Computational Mathematics and Mathematical Physics*, vol. 4, no. 5, pp. 1–17, 1964.

[14] Andrew R Conn, Nicholas IM Gould, and Ph L Toint, *Trust region methods*, vol. 1, Siam, 2000.

[15] David Yang Gao, *Duality principles in nonconvex systems: theory, methods and applications*, vol. 39, Springer Science & Business Media, 2013.

[16] Erik Hons, *Constrained quadratic optimization: Theory and application for wireless communication systems*, Ph.D. thesis, University of Waterloo, 2001.