Optimization Methods and Software

ISSN: 1055-6788 (Print) 1029-4937 (Online) Journal homepage: http://www.tandfonline.com/loi/goms20

Theory and application of p-regularized subproblems for p>2

Yong Hsia, Ruey-Lin Sheu & Ya-xiang Yuan

Optimization Methods & Software, 2017
Vol. 32, No. 5, 1059–1077, https://doi.org/10.1080/10556788.2016.1238917

Theory and application of p-regularized subproblems for $p > 2$

Yong Hsia[a], Ruey-Lin Sheu[b*] and Ya-xiang Yuan[c]

[a]*State Key Laboratory of Software Development Environment, LMIB of the Ministry of Education, School of Mathematics and System Sciences, Beihang University, Beijing 100191, People's Republic of China;* [b]*Department of Mathematics, National Cheng Kung University, Taiwan;* [c]*State Key Laboratory of Scientific/Engineering Computing, Institute of Computational Mathematics and Scientific/Engineering Computing, The Academy of Mathematics and Systems Sciences, Chinese Academy of Sciences, PO Box 2719, Beijing 100080, People's Republic of China*

(*Received 17 May 2016; accepted 16 September 2016*)

The p-regularized subproblem (p-RS) is the key content of a regularization technique in computing a Newton-like step for unconstrained optimization. The idea is to incorporate a local quadratic approximation of the objective function with a weighted regularization term $(\sigma/p)\|x\|^p$ and then globally minimize it at each iteration. In this paper, we establish a complete theory of the p-RSs for general $p > 2$ that covers previous known results on $p = 3$ or $p = 4$. The theory features necessary and sufficient optimality conditions for the global and also for the local non-global minimizers of (p-RS). It gives a closed-form expression for the global minimum set of (p-RS) and shows that (p-RS), $p > 2$ can have at most one local non-global minimizer. Our theory indicates that (p-RS) have all properties that the trust region subproblems do. In application, (p-RS) can appear in natural formulation for optimization problems. We found two examples. One is to utilize the Tikhonov regularization to stabilize the least square solution for an over-determined linear system; and the other comes from numerical approximations to the generalized Ginzburg–Landau functionals. Moreover, when (p-RS) is appended with m additional linear inequality constraints, denoted by (p-RS$_m$), the problem becomes NP-hard. We show that the partition problem, the k-dispersion-sum problem and the quadratic assignment problem in combinatorial optimization can be equivalently formulated as special types of (p-RS$_m$) with $p = 4$. In the end, we develop an algorithm for solving (p-RS$_m$).

Keywords: nonlinear optimization; combinatorial optimization; weighted regularization; trust-region subproblem; extended trust-region subproblem; local non-global minimizer

AMS Subject Classification: 49K30; 90C46; 90C26

1. Introduction

For an unconstrained optimization problem to minimize f over $\mathbb{R}^n$, Newton's method has an attractive local convergence property near a second-order critical point. However, to ensure the global convergence for Newton's method with an analysable computational complexity, it requires modifications to secure a *sufficient* descent in the value of f. This can be guaranteed by trust region methods. The key idea is to compute a trial step by minimizing the second-order Taylor's expansion of f over a trust region ball centred at the current iterate. It leads to the following

*Corresponding author. Email: rsheu@mail.ncku.edu.tw

trust region subproblem:

$$\begin{aligned}\text{(TRS)}\ \min\quad & \tfrac{1}{2}x^{\mathrm{T}}Hx + c^{\mathrm{T}}x \\ \text{s.t.}\quad & \|x\|^2 \le \Delta, \quad x \in \mathbb{R}^n,\end{aligned}$$

where H is the Hessian of f at the current iterate. Because of the compactness of the trust region ball $\{x \mid \|x\|^2 \le \Delta\}$, trust region subproblem (TRS) always has a global minimizer, and it can be solved via a semi-definite program followed by a rank-one decomposition procedure. In particular, we do not need to assume positive definiteness of H. A merit function is then used to determine whether the global minimizer of (TRS) is to be accepted or rejected, followed by an update to the next (TRS). For detailed discussions on trust region methods, see the monograph [5] and the very recent review paper [22].

A relatively new approach is the p-regularized methods. In the subproblems considered by the p-regularized methods, the trust region ball $\{x \mid \|x\|^2 \le \Delta\}$ in (TRS) is replaced with a weighted (by $\sigma > 0$) higher-order regularization term. Specifically, the p-regularized subproblem (p-RS) is the unconstrained minimization problem:

$$(p\text{−RS})\ \min_{x\in\mathbb{R}^n} \left\{ g(x) = \frac{1}{2}x^{\mathrm{T}}Hx + c^{\mathrm{T}}x + \frac{\sigma}{p}\|x\|^p \right\}, \quad p > 2$$

where $\sigma > 0$. Because of the regularization term $(\sigma/p)\|x\|^p$, $g(x)$ is coercive, that is, $\lim_{\|x\|\to+\infty} g(x) = +\infty$ that (p-RS) can always attain the global minimum even for non-positive-definite H. The idea is similar to the trust region algorithm. At any iteration, a local approximation (p-RS) of f is constructed and solved. If the global minimizer of (p-RS) renders a satisfactory decrease in the value of f, it is accepted; but rejected otherwise with an increase in σ to enhance the regularization force. Notice that, due to the regularization term $(\sigma/p)\|x\|^p$, $p > 2$, (p-RS) cannot be formulated and solved by a semi-definite program or by polynomial optimization methods. Normally, it is done by seeking the unique root from a secular equation [10]. Nevertheless, there exist some 'hard' cases [10] in which the computation of a global solution to (p-RS) becomes cumbersome.

In literature, the most common choice to regularize the quadratic approximation is (p-RS) with $p = 3$, which is known as the cubic regularization. The cubic regularization was first introduced in [11] and later was considered by many authors with global convergence and complexity analysis, see [3,18,20]. Recently, a comprehensive comparison for the numerical effectiveness between (p-RS) for general $p > 2$ and (TRS) was made in [10].

Our paper establishes a complete theory of (p-RS) for general $p > 2$ that covers previous known results on $p = 3$ or $p = 4$. The theory includes necessary and sufficient optimality conditions for the global minimum, as well as for the local non-global minimizers, of (p-RS) with $p > 2$. It gives a closed-form expression for the global minimum set, which facilitates the computation of (p-RS) at each step even for the 'hard' case mentioned in [10]. We prove that (p-RS), $p > 2$ can have at most one local non-global minimizer, which enables us to develop an algorithm for solving (p-RS) with additional linear constraints. Our theory shows that (p-RS) have all properties that (TRS) do and thus implies that one can exchange with the other freely. This provides flexibility in formulation and approximation for optimization models. We summarize and comment the main results of the paper as follows.

Theorem 1.1 *The point x^* is a global minimizer of (p-RS) for $p > 2$ if and only if*

$$(H + \sigma\|x^*\|^{p-2}I)x^* = -c; \quad H + \sigma\|x^*\|^{p-2}I \succeq 0. \tag{1}$$

Moreover, the ℓ_2 norms of all the global minimizers are equal.

Theorem 1.1 generalizes a parallel result for $p = 3$ in Theorem 3.1 [3] and complete the sufficient part for $p > 2$ in Theorem 2 [10].

THEOREM 1.2 *Let k be the multiplicity of the smallest eigenvalue α_1 of H, that is ,*

$$\alpha_1 = \cdots = \alpha_k < \alpha_{k+1} \leq \cdots \leq \alpha_n.$$

Then, the set of the global minimizers of (p-RS) is either a singleton or a k-dimensional sphere centred at $(0, \cdots, 0, -c_{k+1}/(\alpha_{k+1} - \alpha_1), \cdots, -c_n/(\alpha_n - \sigma_1))$ with the radius $\sqrt{(\alpha_1/\sigma)^{2/(p-2)} - \sum_{i=k+1}^{n}(c_i^2/(\alpha_i - \alpha_1)^2)}$.

Although (1) in Theorem 1.1 gives the necessary and sufficient condition of the global minimizer for (p-RS), yet the conditions themselves cannot be directly solved to compute the global minimizer of it. It is computed in [10] by finding the root of a secular equation; and in some hard cases with an additional help from eigenvectors and constructed trajectories. Our careful analysis in Theorem 1.2 gives the closed-form expression for the set of global minimizers of (p-RS), $p > 2$. There is no need to distinguish the 'hard case' [10] anymore.

THEOREM 1.3 *The point $\underline{x}$ is a local-non-global minimizer of* (p-RS) *for $p > 2$ if and only if*

$$\underline{x} = -(H + \sigma \underline{t}^* I)^{-1} c,$$

where $\underline{t}^*$ is a root of the secular function

$$h(t) = \|(H + \sigma t I)^{-1} c\|^2 - t^{2/(p-2)}, \quad t \in \left(\max\left\{-\frac{\alpha_2}{\sigma}, 0\right\}, -\frac{\alpha_1}{\sigma}\right) \tag{2}$$

such that $h'(\underline{t}^*) > 0$.

The result is an extension to general $p > 2$ from a special case $p = 4$ in Theorem 1.2 [21]. Moreover, the result is stronger than a parallel version for (TRS) [17] which stated that, if $\underline{x}$ is a local-non-global minimizer of (TRS), then $\underline{x}$ satisfies $(H + \lambda^* I)\underline{x} = -c$ with $\lambda^* \in (-\alpha_2, -\alpha_1)$, $\lambda^* \geq 0$ and $\phi'(\lambda^*) \geq 0$ where $\phi(\lambda) = \|(H + \lambda I)^{-1} c\|^2$. It is not known so far whether or not the necessary condition $\phi'(\lambda^*) \geq 0$ is also sufficient for the local non-global minimizer of (TRS).

THEOREM 1.4 *The subproblem (p-RS) with $p > 2$ has at most one local non-global minimizer. The same property is also shared by (TRS) (proved in [17]) and by the double well potential function ($p = 4$, proved in [21]). The proof for a general theory $p > 2$ here requires special technique to overcome the difficulty.*

The practical applications of (p-RS), however, are not limited to just one of the numerical schemes for nonlinear optimization. It can also appear naturally in formulation for optimization problems. Beck and Ben-Tal [1] utilize the Tikhonov regularization to stabilize the least square solution for an over-determined linear system $Ax = b$ and their model is to solve the quadratic fractional problem

$$\text{(TRTLS)} \quad \min_{x\in\mathbb{R}^n} \left\{ T(x) \equiv \frac{\|Ax-b\|^2}{\|x\|^2+1} + \rho\|x\|^2 \right\},$$

where $\rho > 0$ is a penalty parameter. Then, with Dinkelbach's approach [6], it leads to determine the unique root t^* of a strictly decreasing function:

$$\begin{aligned}\phi(t) &= \min_{x\in\mathbb{R}^n}\{\|Ax-b\|^2 + \rho\|x\|^4 + \rho\|x\|^2 - t(\|x\|^2+1)\} \\ &= \min_{x\in\mathbb{R}^n}\{x^{\mathrm{T}}A^{\mathrm{T}}Ax + (\rho - t)\|x\|^2 - 2b^{\mathrm{T}}Ax + \rho\|x\|^4\} + \|b\|^2 - t. \end{aligned} \tag{3}$$

In each step, Algorithm TRTLSI [1] evaluates $\phi(t)$ and updates t until the sequence converges to t^* eventually. Notice that, for each t, evaluating $\phi(t)$ in (3) amounts to solving (p-RS) with $p=4$. Here we point out that formula (54) given in [1] is not correct. The correct formula (please refer to notations [1]) should be

$$\begin{aligned}\phi(t) = \min_{z\in\mathbb{R}^n} &\left\{ \sum_{j=1}^{n} \left\{ \lambda_j z_j^2 + (\rho - t)z_j^2 - 2f_j z_j + \rho z_j^4 + 2\rho \sum_{i<j} z_i^2 z_j^2 \right\}\right\} \\ &+ \|b\|^2 - t. \end{aligned} \tag{4}$$

Therefore, all subsequent analysis and related computation after (54) [1], including Algorithm TRTLSI itself, should be modified accordingly. Given that the correct formula (4) is more complicate than the wrong one (54) used [1], we suggest that our result in this paper can be directly incorporated with Algorithm TRTLSI to simplify the implementation.

Another practical application of (p-RS) comes from numerical approximations to the generalized Ginzburg–Landau functionals [16]:

$$I^{\alpha}(\mu) = \int_{\Omega} \left[\frac{1}{n}\|\nabla\mu(x)\|^n + \frac{\alpha}{2}\left(\frac{1}{2}\|\mu(x)\|^2 - \beta\right)^2 \right] \mathrm{d}x, \tag{5}$$

where $\Omega \subset R^n$, α, β are positive material constants, and $\mu : \Omega \longrightarrow R^q$ is a smooth vector-valued (field) function describing the phase of the system. The second term of (5), $\int_{\Omega}(\alpha/2)(\frac{1}{2}\|\mu(x)\|^2 - \beta)^2]\,\mathrm{d}x$, is called the double-well potential in the integral form. Discretizing it leads naturally to a form of (p-RS) with $p=4$. Please refer to [8] for detail derivation.

The parallel structure between (TRS) and (p-RS) extends to some recent results in cases when additional m linear inequality constraints are added:

$$(p\text{-RS}_m) \quad \min \frac{1}{2}x^{\mathrm{T}}Hx + c^{\mathrm{T}}x + \frac{\sigma}{p}\|x\|^p \tag{6}$$

$$\text{s.t. } l_i \le a_i{}^{\mathrm{T}}x \le u_i, \quad i = 1, \ldots, m, \tag{7}$$

where $l_i \le u_i \in \mathbb{R}$ for $i = 1, \ldots, m$. When (TRS) is appended with linear inequality constraints, it is called the extended trust region subproblem (ETRS). Polynomial solvability for (ETRS) when the number m is fixed has been recently proved in [2], and independently in [14]. Both methods enumerate all the intersecting faces, while recursively reducing the problem dimension as well as the number of constraints until it has just one linear constraint remained. The enumeration is exponential in the number m, as it contributes to the combinatorial nature on the boundary of a polytope. When m is fixed, the algorithm has a polynomial complexity with respect to all other data size of the problem. We refer the reader to [2,14] for more references regarding (ETRS) and the reduction algorithms.

In the remaining part of the paper, we demonstrate that (p-RS$_m$) can be also used to reformulate combinatorial optimization problems; and it happens that the reduction method for (ETRS) [14] has a non-trivial resemblance for solving (p-RS$_m$) with the same complexity, particularly for $p = 4$. We want to emphasize that, due to the regularization term $(\sigma/p)\|x\|^p$, Bienstock and Michalka's approach [2] does not inherently entitle a direct extension to solve (p-RS$_m$) here.

For (p-RS$_m$) to formulate combinatorial optimization problems, we give two examples. By the first one we show that the NP-hard partition problem, which checks, for any given positive integer vector r, the solvability of

$$r^{\mathrm{T}}x = 0, \quad x \in \{-1, 1\}^n,$$

can be reduced to a special case of (p-RS$_n$). Secondly, we show that the NP-hard binary quadratic optimization problem

$$\text{(BQP)} \quad \min\ x^{\mathrm{T}}Qx$$
$$\text{s. t. } e^{\mathrm{T}}x = k, \quad Ax \le b, \quad x \in \{0, 1\}^n$$

is reduced to a special case of (p-RS$_{m+n+1}$) with $p = 4$. The binary quadratic program (BQP) includes the k-dispersion-sum problem and the quadratic assignment problem as special cases. Both the partition problem and the binary quadratic optimization problem can therefore be solved by the reduction algorithm that we develop for (p-RS$_m$).

In the concluding remark of the paper, we mention that the free interchange between the trust region constraint $\|x\|^2 \le \Delta$ in (TRS) and the weighted regularization term $(\sigma/p)\|x\|^p$ in (p-RS) can have meaningful implications. Especially for the Celis-Dennis-Tapia (CDT) problem [4], their formulation requires to solve a quadratic approximation on the intersection of one ball and one ellipsoid, which truly introduces enormous difficulty. If (TRS) is replaced with (p-RS) in the CDT formulation, an obvious advantage is that we only have to consider (p-RS) subject to just one ellipsoid constraint. That will be an interesting future research topic.

Notation 1 Let $v(\cdot)$ denote the optimal value of problem $(\cdot)$. For any symmetric matrix $P \in \mathbb{R}^{n\times n}$, $P \succ (\succeq) 0$ means that P is positive (semi)definite. The determinant of P is denoted by $\det(P)$ whereas the identity matrix of order n by I. For a vector $x \in \mathbb{R}^n$, Diag(x) is a diagonal matrix with diagonal components being $x_1, \ldots, x_n$. e denotes a vector of dimension n with all components equal to one. For a number $\beta \in \mathbb{R}$, $\mathrm{sign}(\beta) = \beta/|\beta|$ if $\beta \neq 0$, otherwise $\mathrm{sign}(\beta) = 0$. Finally, $\lambda_i(P)$ is the i-th smallest eigenvalue of P.

2. Characterization of the global minimizers

The starting point of the analysis is the first-order and the second-order necessary conditions for any local minimizer of g.

LEMMA 2.1 *Assume that $\underline{x}$ is a local minimizer of (p-RS), $p > 2$. It holds that*

$$\nabla g(\underline{x}) = (H + \sigma\|\underline{x}\|^{p-2}I)\underline{x} + c = 0, \tag{8}$$

$$\nabla^2 g(\underline{x}) = (H + \sigma\|\underline{x}\|^{p-2}I) + \sigma(p-2)\|\underline{x}\|^{p-4}\underline{x}\underline{x}^{\mathrm{T}} \succeq 0, \tag{9}$$

where ∇g, $\nabla^2 g$ denote the gradient and the Hessian of $g(x)$, respectively.

The next theorem shows that, a local minimizer $\underline{x}$ becomes global if and only if $H + \sigma\|\underline{x}\|^{p-2}I \succeq 0$. The necessity has been shown by Theorem 1.2 [10]. We only prove the sufficiency here.

THEOREM 2.2 *The point x^* is a global minimizer of (p-RS) for $p > 2$ if and only if it is a critical point satisfying $\nabla g(x^*) = 0$ and $H + \sigma\|x^*\|^{p-2}I \succeq 0$. Moreover, the ℓ_2 norms of all the global minimizers are equal.*

Proof If $x^* = 0_n$, then $\sigma\|x^*\|^{p-2} = 0$ so that $c = -(H + \sigma\|x^*\|^{p-2}I)x^* = 0$ and $H = H + \sigma\|x^*\|^{p-2}I \succeq 0$. Consequently, $x^{\mathrm{T}}Hx \geq 0,\ \forall x \in \mathbb{R}^n$. It follows that $x^* = 0_n$ is a global minimizer since

$$g(x) = \frac{1}{2}x^{\mathrm{T}}Hx + c^{\mathrm{T}}x + \frac{\sigma}{p}\|x\|^p \geq \frac{\sigma}{p}\|x\|^p > 0 = g(0), \quad \forall x \neq 0_n = x^*.$$

Now we assume $x^* \neq 0_n$, that is, $\|x^*\| > 0$. Define $Q = H + \sigma\|x^*\|^{p-2}I$. According to the assumption, $Q \succeq 0$. Then, for any $x \in \mathbb{R}^n$ and $x \neq x^*$, it holds that

$$\begin{aligned} g(x) &= \frac{1}{2}x^{\mathrm{T}}Hx + c^{\mathrm{T}}x + \frac{\sigma}{p}\|x\|^p \\ &= \frac{1}{2}x^{\mathrm{T}}Qx + c^{\mathrm{T}}x - \frac{1}{2}(\sigma\|x^*\|^{p-2})x^{\mathrm{T}}x + \frac{\sigma}{p}\|x\|^p \\ &= \frac{1}{2}x^{\mathrm{T}}Qx + c^{\mathrm{T}}x + \frac{\sigma}{p}\|x^*\|^p\left(\left(\frac{\|x\|^2}{\|x^*\|^2}\right)^{\frac{p}{2}} - \frac{p}{2}\frac{\|x\|^2}{\|x^*\|^2}\right). \end{aligned} \tag{10}$$

Define $f(t) = t^{p/2},\ p > 2$. It is strictly convex for $t > 0$. Therefore,

$$f(t) = t^{p/2} \geq f(1) + f'(1)(t-1) = 1 + \frac{p}{2}(t-1), \quad \forall t > 0.$$

By substituting t with $\frac{\|x\|^2}{\|x^*\|^2}$, we have

$$\left(\frac{\|x\|^2}{\|x^*\|^2}\right)^{\frac{p}{2}} - \frac{p}{2}\frac{\|x\|^2}{\|x^*\|^2} \geq 1 - \frac{p}{2}.$$

Then,

$$g(x) \geq \frac{1}{2}x^{\mathrm{T}}Qx + c^{\mathrm{T}}x + \frac{\sigma}{p}\|x^*\|^p\left(1 - \frac{p}{2}\right). \tag{11}$$

By $Q \succeq 0$, the lower bounding function of g in the right-hand side of (11) is convex quadratic in terms of x. Since x^* satisfies $(H + \sigma\|x^*\|^{p-2}I)x^* = Qx^* = -c$, x^* is a global minimizer of the convex function in the right-hand side of (11). As a consequence,

$$g(x) \geq \frac{1}{2}(x^*)^{\mathrm{T}}Qx^* + c^{\mathrm{T}}x^* + \frac{\sigma}{p}\|x^*\|^p\left(1 - \frac{p}{2}\right) = g(x^*)$$

and x^* is a global minimizer of (p-RS).

Finally, from (10), if $\hat{x}$ is also a global minimizer of (p-RS), $\hat{x}$ must minimize both $\frac{1}{2}x^{\mathrm{T}}Qx + c^{\mathrm{T}}x$ and $(\|x\|^2/\|x^*\|^2)^{p/2} - (p/2)(\|x\|^2/\|x^*\|^2)$ simultaneously since x^* does too. This can happen if and only if $Q\hat{x} = -c$ and $\|\hat{x}\| = \|x^*\|$. ■

Remark 1 When $p = 3$, two other proofs of the necessary and sufficient condition can be found in Theorem 3.1 [3] and Theorem 10 [18], respectively. We notice that the proof in [3] is inherited from that of the necessary and sufficient condition for the trust-region subproblem [5] and the proof in [18] highly relies on the special structure of the case $p = 3$. Our proof is much easier to understand, since it is based on a direct comparison between $g(x)$ and a convex lower bound function.

To characterize the set of global minimizers of (p-RS), we may assume that H is diagonal, that is,

$$H = \mathrm{Diag}(\alpha_1, \ldots, \alpha_n), \tag{12}$$

where

$$\alpha_1 = \cdots = \alpha_k < \alpha_{k+1} \le \cdots \le \alpha_n$$

and k is the multiplicity of the smallest eigenvalue α_1. Otherwise, let $H = U\Sigma U^{\mathrm{T}}$ be the eigenvalue decomposition of H. Let $y = U^{\mathrm{T}}x$. Then $\|y\| = \|U^{\mathrm{T}}x\| = \|x\|$ and we obtain a diagonal version of (p-RS) in terms of y.

THEOREM 2.3 *The set of global minimizers of* (p-RS) *is either a singleton or a k-dimensional sphere centred at* $(0, \ldots, 0, -c_{k+1}/(\alpha_{k+1} - \alpha_1), \ldots, -c_n/(\alpha_n - \alpha_1))$ *with the radius* $\sqrt{(\alpha_1/\sigma)^{2/(p-2)} - \sum_{i=k+1}^n (c_i^2/(\alpha_i - \alpha_1)^2)}$.

Proof Let x^* be any global minimizer of (p-RS) and define $t^* = \|x^*\|^{p-2} \ge 0$. Notice that t^* is independent of the choice of x^* since the ℓ_2 norms of all the global minimizers are equal. By Theorem 2.2, $\alpha_i + \sigma t^* \ge 0,\ \forall i = 1, 2, \ldots, n$. If $H + \sigma t^* I$ is invertible, $t^* \in (\max\{-\alpha_1/\sigma, 0\}, +\infty)$ and the global minimizer x^* is uniquely defined by (the still unknown t^* that)

$$x_i^* = \frac{-c_i}{\alpha_i + \sigma t^*}, \quad i = 1, \ldots, n.$$

By summing all $(x_i^*)^2$, t^* is necessarily a non-negative root of the following secular function on a specific open interval:

$$h(t) = \sum_{i=1}^n \frac{c_i^2}{(\alpha_i + \sigma t)^2} - t^{\frac{2}{p-2}}, \quad t \in I_g = \left(\max\left\{-\frac{\alpha_1}{\sigma}, 0\right\}, +\infty\right). \tag{13}$$

Since $\lim_{t\to\max\{-\alpha_1/\sigma,0\}} h(t) > 0$, $\lim_{t\to+\infty} h(t) = -\infty$ and $h(t)$ is strictly decreasing on I_g (see Remark 2), the secular function $h(t)$ has a unique root on I_g, which must be t^*.

On the other hand, $H + \sigma t^* I$ is singular in which case $t^* = -\alpha_1/\sigma$. (Obviously, this case cannot happen for $\alpha_1 > 0$.) Then, $c_1^2 + \cdots + c_k^2 = 0$, and $\alpha_i + \sigma t^* > 0,\ i = k+1, k+2, \ldots, n$ such that

$$\hat{x}^* = \left(0, 0, \ldots, 0, \frac{-c_{k+1}}{\alpha_{k+1} - \alpha_1}, \ldots, \frac{-c_n}{\alpha_n - \alpha_1}\right)^{\mathrm{T}} \tag{14}$$

is one trivial solution to $(H - \alpha_1 I)x^* = -c$. By summing all $(\hat{x}_i^*)^2$ in (14), we again obtain a secular function

$$\hat{h}(t) = \sum_{i=k+1}^n \frac{c_i^2}{(\alpha_i + \sigma t)^2} - t^{2/(p-2)}, \quad t \in I_{\hat{g}} = \left[-\frac{\alpha_1}{\sigma}, +\infty\right). \tag{15}$$

Notice that $\hat{h}(t)$ is also strictly decreasing on $I_{\hat{g}}$ and $\lim_{t\to+\infty} \hat{h}(t) = -\infty$. If $\hat{h}(-\alpha_1/\sigma) = 0$, then $t^* = -\alpha_1/\sigma$ is the unique root of $\hat{h}(t)$ on $I_{\hat{g}}$. Thus, $\hat{x}^*$ defined by (14) is the unique global minimizer of (p-RS).

If $\hat{h}(-\alpha_1/\sigma) < 0$, then (15) has no solution and the trivial solution $\hat{x}^*$ to $(H - \alpha_1 I)x^* = -c$ does not satisfy $t^* = -\alpha_1/\sigma = \|\hat{x}^*\|^{p-2}$. Then, any x^* satisfying

$$(x_1^*)^2 + \cdots + (x_k^*)^2 + \sum_{i=k+1}^{n} \frac{c_i^2}{(\alpha_i - \alpha_1)^2} = \left(\frac{-\alpha_1}{\sigma}\right)^{\frac{2}{p-2}} \tag{16}$$

is a global minimizer of (p-RS). Namely, the global minimum solution set forms a k-dimensional sphere centred at $(0, \ldots, 0, -c_{k+1}/(\alpha_{k+1} - \alpha_1), \ldots, -c_n/(\alpha_n - \alpha_1))$ with the radius $\sqrt{(\alpha_1/\sigma)^{2/(p-2)} - \sum_{i=k+1}^{n}(c_i^2/(\alpha_i - \alpha_1)^2)}$.

Otherwise, $\hat{h}(-\alpha_1/\sigma) > 0$, then (15) has no solution and (16) cannot hold for any x^*. We obtain a contradiction that (p-RS) has no global minimizer. ■

Finally in this section, we show that (p-RS) possesses some hidden convexity that its global minimizer can be obtained by solving an equivalently reformulated convex programming. We first have

Proposition 2.4 *Suppose H is diagonal. Let x^* be any global minimizer of* (p-RS), *then*

$$c_i x_i^* \le 0, \quad i = 1, \ldots, n.$$

Proof Comparing x^* with $\tilde{x} = (-x_1^*, x_2^*, x_3^*, \ldots, x_n^*)$, we immediately have

$$0 \ge g(x^*) - g(\tilde{x}) = c_1(x_1^* - \tilde{x}_1) = 2c_1 x_1^*.$$

A similar argument applying to all other components yields the result. ■

By Proposition 2.4, (p-RS) and (17) below share the same optimal solution set.

$$\begin{aligned} \min \quad & \sum_{i=1}^{n} \left\{ \frac{\alpha_i}{2} x_i^2 + c_i x_i \right\} + \frac{\sigma}{p} \left(\sum_{i=1}^{n} x_i^2 \right)^{\frac{p}{2}} \\ \text{s.t.} \quad & c_i x_i \le 0, \quad i = 1, \ldots, n. \end{aligned} \tag{17}$$

Introducing the nonlinear one-to-one map:

$$x_i = \begin{cases} \sqrt{z_i}, & \text{if } c_i \le 0, \\ -\sqrt{z_i}, & \text{if } c_i > 0, \end{cases} \quad i = 1, \ldots, n \tag{18}$$

the problem (17) becomes the following convex program:

$$\begin{aligned} \min \quad & -\sum_{i=1}^{n} |c_i| \sqrt{z_i} + \frac{1}{2} \sum_{i=1}^{n} \alpha_i z_i + \frac{\sigma}{p} \left(\sum_{i=1}^{n} z_i \right)^{\frac{p}{2}} \\ \text{s.t.} \quad & z_i \ge 0, \quad i = 1, \ldots, n. \end{aligned} \tag{19}$$

The global optimal solution of (19) can be converted to generate x^* through the transformation (18).

Remark 2 The first two derivatives of the secular function $h(t)$ are

$$h'(t) = \sum_{i=1}^{n} \frac{-2\sigma c_i^2}{(\alpha_i + \sigma t)^3} - \frac{2}{p-2} t^{\frac{4-p}{p-2}}$$

and

$$h''(t) = \sum_{i=1}^{n} \frac{6\sigma^2 c_i^2}{(\alpha_i + \sigma t)^4} - \frac{2(4-p)}{(p-2)^2}\, t^{\frac{6-2p}{p-2}}.$$

Note that $h(t)$ is strictly decreasing on I_g and convex only for $p \geq 4$. For $p = 3$, $h(t)$ can be made convex by properly choosing the regularization parameter σ when $H \nsucceq 0$ (which ensures that $c \neq 0$) and h is restricted to a finite subinterval of I_g covering t^*. Nevertheless, the secular function for (TRS) is always convex.

3. Characterization of the local-non-global minimizer

Throughout this section, we assume that H is diagonal as in (12) and $\alpha_1 < 0$. If $\alpha_1 \geq 0$, (p-RS) is a convex minimization problem and hence has no local-non-global minimum. Moreover, 0_n cannot be a local non-global minimizer since the necessary optimality conditions (8)– (9) at 0_n imply that $c = 0$, $H \succeq 0$. In characterizing the local non-global minimum we need to use the second smallest eigenvalue α_2 of H, so we implicitly assume $n \geq 2$. For $n = 1$, it can be treated as if $\alpha_2 = \infty$ in each of the related theorems.

LEMMA 3.1 *Suppose $\underline{x}$ is a local non-global minimizer of (p-RS). It holds that $\underline{x}_1 \neq 0$, $\alpha_1 < \alpha_2$ and*

$$\alpha_2 + \sigma \|\underline{x}\|^{p-2} > 0. \tag{20}$$

Proof Since $\underline{x}$ is a local but non-global minimizer, by Theorem 2.2, $H + \sigma \|\underline{x}\|^{p-2} I \nsucceq 0$. It follows immediately that $\alpha_1 + \sigma \|\underline{x}\|^{p-2} < 0$. By the second-order condition for $\underline{x}$ in (9), the first two leading principal submatrices of $\nabla^2 g(\underline{x})$ are positive semidefinite. Namely,

$$\alpha_1 + \sigma \|\underline{x}\|^{p-2} + \sigma (p-2) \|\underline{x}\|^{p-4} \underline{x}_1^2 \geq 0 \tag{21}$$

and

$$\begin{bmatrix} \alpha_1 + \sigma \|\underline{x}\|^{p-2} & 0 \\ 0 & \alpha_2 + \sigma \|\underline{x}\|^{p-2} \end{bmatrix} + \sigma (p-2) \|\underline{x}\|^{p-4} \begin{bmatrix} \underline{x}_1^2 & \underline{x}_1 \underline{x}_2 \\ \underline{x}_1 \underline{x}_2 & \underline{x}_2^2 \end{bmatrix} \succeq 0. \tag{22}$$

Since $\alpha_1 + \sigma \|\underline{x}\|^{p-2} < 0$, it follows from (21) that $\underline{x}_1 \neq 0$. Moreover, since

$$(-\underline{x}_2, \underline{x}_1) \begin{bmatrix} \underline{x}_1^2 & \underline{x}_1 \underline{x}_2 \\ \underline{x}_1 \underline{x}_2 & \underline{x}_2^2 \end{bmatrix} \begin{pmatrix} -\underline{x}_2 \\ \underline{x}_1 \end{pmatrix} = 0,$$

we have from (22) that

$$(-\underline{x}_2, \underline{x}_1) \begin{bmatrix} \alpha_1 + \sigma \|\underline{x}\|^{p-2} & 0 \\ 0 & \alpha_2 + \sigma \|\underline{x}\|^{p-2} \end{bmatrix} \begin{pmatrix} -\underline{x}_2 \\ \underline{x}_1 \end{pmatrix}$$
$$= (\alpha_1 + \sigma \|\underline{x}\|^{p-2})(\underline{x}_2)^2 + (\alpha_2 + \sigma \|\underline{x}\|^{p-2})(\underline{x}_1)^2 \geq 0.$$

As $(\alpha_1 + \sigma \|\underline{x}\|^{p-2})(\underline{x}_2)^2 \leq 0$ and $\underline{x}_1 \neq 0$, it holds that $\alpha_2 + \sigma \|\underline{x}\|^{p-2} \geq 0$. Due to $\alpha_1 + \sigma \|\underline{x}\|^{p-2} < 0$, we know $\alpha_1 \neq \alpha_2$.

To argue that $\alpha_2 + \sigma\|\underline{x}\|^{p-2} \neq 0$, we assume the contrary that $\alpha_2 + \sigma\|\underline{x}\|^{p-2} = 0$ and show a contradiction. In this case, the second-order necessary condition (9) reduces to

$$H - \alpha_2 I + \sigma(p-2)\|\underline{x}\|^{p-4}\underline{x}\underline{x}^{\mathrm{T}} \succeq 0.$$

By $\alpha_1 - \alpha_2 < 0$ and

$$\begin{aligned}\det&\left\{\begin{bmatrix}\alpha_1-\alpha_2 & 0\\ 0 & 0\end{bmatrix} + \sigma(p-2)\|\underline{x}\|^{p-4}\begin{bmatrix}\underline{x}_1^2 & \underline{x}_1\underline{x}_2\\ \underline{x}_1\underline{x}_2 & \underline{x}_2^2\end{bmatrix}\right\}\\ &= \sigma(p-2)\|\underline{x}\|^{p-4}(\alpha_1-\alpha_2)\underline{x}_2^2 \geq 0,\end{aligned}$$

it implies that $\underline{x}_2 = 0$ and, from the first-order necessary condition (8),

$$\underline{x}_1 = \frac{-c_1}{\alpha_1 + \sigma\|\underline{x}\|^{p-2}} = \frac{c_1}{\alpha_2 - \alpha_1}.$$

Since $\underline{x}_1 \neq 0$, therefore $c_1 \neq 0$. Without loss of generality, we may assume both $c_1 > 0$ and $\underline{x}_1 > 0$. Define

$$k(t) = \sqrt{\underline{x}_1^2 - t^2}, \quad t \in [-\underline{x}_1, \underline{x}_1]$$

and consider the following parametric curve in $\mathbb{R}^n$:

$$\gamma(t) = \{(k(t), t, \underline{x}_3, \ldots, \underline{x}_n) | t \in [-\underline{x}_1, \underline{x}_1]\}. \tag{23}$$

Notice that $\gamma(0) = \gamma(\underline{x}_2) = \underline{x}$, that is, $\gamma(t)$ passes through $\underline{x}$ at $t=0$. Evaluating $g(x)$ on $\gamma(t)$, we have

$$\begin{aligned}&g(\gamma(t))\\ &= \frac{\sigma}{p}\left(k(t)^2 + t^2 + \sum_{i=3}^{n}\underline{x}_i^2\right)^{\frac{p}{2}} + \frac{\alpha_1}{2}k(t)^2 + \frac{\alpha_2}{2}t^2 + \sum_{i=3}^{n}\frac{\alpha_i}{2}\underline{x}_i^2 + c_1 k(t) + \sum_{i=3}^{n}c_i\underline{x}_i\\ &= \frac{\sigma}{p}\left(\underline{x}_1^2 + \sum_{i=3}^{n}\underline{x}_i^2\right)^{\frac{p}{2}} + \frac{\alpha_1}{2}\underline{x}_1^2 + \sum_{i=3}^{n}\frac{\alpha_i}{2}\underline{x}_i^2 + \frac{\alpha_2-\alpha_1}{2}t^2 + c_1\sqrt{\underline{x}_1^2 - t^2} + \sum_{i=3}^{n}c_i\underline{x}_i.\end{aligned}$$

Since $\underline{x}$ is a local minimizer of $g(x)$, $t=0$ must be a local minimum point of $g(\gamma(t))$. However, this implication contradicts to the fact that

$$\frac{\mathrm{d}}{\mathrm{d}t}g(\gamma(0)) = \frac{d^2}{\mathrm{d}t^2}g(\gamma(0)) = \frac{d^3}{\mathrm{d}t^3}g(\gamma(0)) = 0, \frac{d^4}{\mathrm{d}t^4}g(\gamma(0)) = -\frac{3(\alpha_2-\alpha_1)}{\underline{x}_1^2} < 0.$$

■

The following theorem on necessary and sufficient conditions for any local-non-global minimizer of (p-RS) is the main result in this section.

THEOREM 3.2 *$\underline{x}$ is a local-non-global minimizer of (p-RS) if and only if*

$$\underline{x} = -(H + \sigma\underline{t}^* I)^{-1}c, \tag{24}$$

where $\underline{t}^*$ is a root of the secular function

$$h(t) = \|(H + \sigma t I)^{-1}c\|^2 - t^{2/(p-2)}, \quad t \in \left(\max\left\{-\frac{\alpha_2}{\sigma}, 0\right\}, -\frac{\alpha_1}{\sigma}\right) \tag{25}$$

such that $h'(\underline{t}^*) > 0$.

Proof According to Lemma 3.1 and Theorem 2.2, the local-non-global minimizer $\underline{x}$ of (p-RS) exits only if $\alpha_1 < \alpha_2$ and

$$-\alpha_2 < \sigma \|\underline{x}\|^{p-2} < -\alpha_1.$$

Then, the diagonal matrix $H + \sigma\|\underline{x}\|^{p-2}I$ is non-singular and the first-order necessary condition (8) can be solved by

$$\underline{x}_i = \frac{-c_i}{\alpha_i + \sigma\|\underline{x}\|^{p-2}}, \quad i = 1, \ldots, n \tag{26}$$

and therefore

$$\|\underline{x}\|^2 = \sum_{i=1}^{n} \frac{c_i^2}{(\alpha_i + \sigma\|\underline{x}\|^{p-2})^2}.$$

Define $\underline{t}^* = \|\underline{x}\|^{p-2}$. It is obvious that $\underline{t}^*$ is a root of the following secular function on a specific open interval:

$$h(t) = \sum_{i=1}^{n} \frac{c_i^2}{(\alpha_i + \sigma t)^2} - t^{\frac{2}{p-2}}, \quad t \in \left(\max\left\{-\frac{\alpha_2}{\sigma}, 0\right\}, -\frac{\alpha_1}{\sigma}\right). \tag{27}$$

Due to the non-singularity of $H + \sigma\underline{t}^*I$, we see from (26) that each local non-global minimizer $\underline{x}$ can be represented by $-(H + \sigma\underline{t}^*I)^{-1}c$, which is (24).

Taking a simple calculation on $h(t)$, we have

$$h'(t) = -\sum_{i=1}^{n} \frac{2\sigma c_i^2}{(\alpha_i + \sigma t)^3} - \frac{2}{p-2} t^{\frac{4-p}{p-2}}.$$

We notice that the necessary optimality condition (9) is equivalent to

$$\mathrm{Diag}(-1, 1, \ldots, 1) + \sigma(p-2)\|\underline{x}\|^{p-4}(\Gamma\underline{x})(\Gamma\underline{x})^{\mathrm{T}} \succeq 0,$$

where

$$\Gamma = \mathrm{Diag}\left(\frac{1}{\sqrt{-\alpha_1 - \sigma\|\underline{x}\|^{p-2}}}, \frac{1}{\sqrt{\alpha_2 + \sigma\|\underline{x}\|^{p-2}}}, \ldots, \frac{1}{\sqrt{\alpha_n + \sigma\|\underline{x}\|^{p-2}}}\right).$$

Then,

$$\begin{aligned}
0 &\le \det(\mathrm{Diag}(-1, 1, \ldots, 1) + \sigma(p-2)\|\underline{x}\|^{p-4}(\Gamma\underline{x})(\Gamma\underline{x})^{\mathrm{T}}) \\
&= \det(\mathrm{Diag}(-1, 1, \ldots, 1)) \\
&\quad \times \det(\sigma(p-2)\|\underline{x}\|^{p-4}\mathrm{Diag}(-1, 1, \ldots, 1)(\Gamma\underline{x})(\Gamma\underline{x})^{\mathrm{T}} + I) \\
&= -1 \times (\sigma(p-2)\|\underline{x}\|^{p-4}(\Gamma\underline{x})^{\mathrm{T}}\mathrm{Diag}(-1, 1, \ldots, 1)(\Gamma\underline{x}) + 1) \\
&= -\sum_{i=1}^{n} \frac{\sigma(p-2)\|\underline{x}\|^{p-4}c_i^2}{(\sigma\|\underline{x}\|^{p-2} + \alpha_i)^3} - 1 \\
&= \left(\frac{p}{2} - 1\right)\|\underline{x}\|^{p-4}h'(\|\underline{x}\|^{p-2}) \\
&= \left(\frac{p}{2} - 1\right)\|\underline{x}\|^{p-4}h'(\underline{t}^*).
\end{aligned}$$

It follows from $p > 2$ and the fact that 0_n cannot be a local non-global minimizer, that $h'(\underline{t}^*) \geq 0$. To show that $h'(\underline{t}^*) > 0$, we assume the contrary that $h'(\underline{t}^*) = 0$. Then,

$$\det(H + \sigma\|\underline{x}\|^{p-2}I + \sigma(p-2)\|\underline{x}\|^{p-4}\underline{x}\underline{x}^{\mathrm{T}}) = \frac{(\frac{p}{2}-1)\|\underline{x}\|^{p-4}h'(\underline{t}^*)}{\det^2(\Gamma)} = 0 \tag{28}$$

and thus there is a $u = (u_1, \ldots, u_n)^{\mathrm{T}} \neq 0$ such that

$$(H + \sigma\|\underline{x}\|^{p-2}I)u + \sigma(p-2)\|\underline{x}\|^{p-4}\underline{x}\underline{x}^{\mathrm{T}}u = 0,$$

or equivalently,

$$u_i = \frac{-\sigma(p-2)\|\underline{x}\|^{p-4}\underline{x}_i(u^{\mathrm{T}}\underline{x})}{\alpha_i + \sigma\|\underline{x}\|^{p-2}}, \quad i = 1, 2, \ldots, n.$$

Since $u \neq 0$, it holds that

$$u^{\mathrm{T}}\underline{x} \neq 0. \tag{29}$$

Define

$$q(\beta) := g(\underline{x} + \beta u).$$

We can verify that

$$\begin{aligned} q'(\beta) &= u^{\mathrm{T}}\nabla g(\underline{x} + \beta u), \\ q''(\beta) &= u^{\mathrm{T}}\nabla^2 g(\underline{x} + \beta u)u, \\ q'''(\beta) &= 3\sigma(p-2)\|\underline{x} + \beta u\|^{p-4}(u^{\mathrm{T}}\underline{x} + \beta u^{\mathrm{T}}u)u^{\mathrm{T}}u \\ &\quad + \sigma(p-2)(p-4)\|\underline{x} + \beta u\|^{p-6}(u^{\mathrm{T}}\underline{x} + \beta u^{\mathrm{T}}u)^3. \end{aligned}$$

The necessary optimality condition (8) implies that $q'(0) = 0$. According to the definition of u, we have $q''(0) = 0$. However, (29) implies that

$$\begin{aligned} (q'''(0))^2 &= \sigma^2(p-2)^2\|\underline{x}\|^{2(p-6)}(u^{\mathrm{T}}\underline{x})^6\left(3\frac{\|\underline{x}\|^2\|u\|^2}{(u^{\mathrm{T}}\underline{x})^2} + (p-4)\right)^2 \\ &\geq \sigma^2(p-2)^2\|\underline{x}\|^{2(p-6)}(u^{\mathrm{T}}\underline{x})^6(p-1)^2 \\ &> 0, \end{aligned}$$

which contradicts to the fact that $\underline{x}$ is a local minimizer of (p-RS). Therefore, $h'(\underline{t}^*) > 0$ and the necessary proof is complete.

For the sufficient part, let $\underline{t}^* \in (\max\{-\alpha_2/\sigma, 0\}, -\alpha_1/\sigma)$ be a root of the secular function (27) such that $h'(\underline{t}^*) > 0$. Define $\underline{x}$ as in (24). Then we have

$$\|\underline{x}\|^2 = \sum_{i=1}^{n} \frac{c_i^2}{(\sigma\underline{t}^* + \alpha_i)^2} = (\underline{t}^*)^{\frac{2}{p-2}}.$$

That is, $\underline{t}^* = \|\underline{x}\|^{p-2}$. Consequently, $\underline{x}$ satisfies the first-order necessary optimality condition (8). Moreover, the diagonal matrix $H + \sigma\|\underline{x}\|^{p-2}I$ is non-singular with positive diagonal elements except for the first one. By Weyl's inequality (see [13], Theorem 4.3.1), we have

$$\begin{aligned}\lambda_i(\nabla^2 g(\underline{x})) &= \lambda_i(H + \sigma\|\underline{x}\|^{p-2}I + \sigma(p-2)\|\underline{x}\|^{p-4}\underline{x}\underline{x}^{\mathrm{T}})\\ &\geq \lambda_i(H + \sigma\|\underline{x}\|^{p-2}I) + \lambda_1(\sigma(p-2)\|\underline{x}\|^{p-4}\underline{x}\underline{x}^{\mathrm{T}})\\ &\geq \lambda_i(H + \sigma\|\underline{x}\|^{p-2}I)\\ &> 0, \quad \text{for } i = 2,3,\ldots,n.\end{aligned}$$

From (28), since $h'(\underline{t}^*) > 0$, we have

$$\begin{aligned}\prod_{i=1}^{n}\lambda_i(\nabla^2 g(\underline{x})) &= \det(\nabla^2 g(\underline{x}))\\ &= \frac{(\frac{p}{2}-1)\|\underline{x}\|^{p-4}h'(\underline{t}^*)}{\det^2(\Gamma)} > 0,\end{aligned}$$

which implies that $\lambda_1(\nabla^2 g(\underline{x})) > 0$ and thus $\nabla^2 g(\underline{x}) \succ 0$. This guarantees that $\underline{x}$ is a local non-global minimizer of (p-RS). The proof is complete. ∎

Remark 1 Our secular function (25) is not new here. Actually, it belongs to the family of secular functions defined in [10]. However, the analysis here is novel.

THEOREM 3.3 *(p-RS) with $p > 2$ has at most one local-non-global minimizer.*

Proof We first observe that the secular function (25) has the same roots as

$$p(t) = \log(\|(H+\sigma tI)^{-1}c\|^2) - \frac{2}{p-2}\log(t), \quad t \in \left(\max\left\{-\frac{\alpha_2}{\sigma}, 0\right\}, -\frac{\alpha_1}{\sigma}\right).$$

By Lemma 3.1 and (26), $\underline{x}_1 \neq 0$ and $c_1 \neq 0$. Then, we have

$$p''(t) = \frac{\sum_{i=1}^{n}\frac{6\sigma^2 c_i^2}{(\alpha_i+\sigma t)^4}}{\sum_{i=1}^{n}\frac{c_i^2}{(\alpha_i+\sigma t)^2}} - \frac{\left(\sum_{i=1}^{n}\frac{2\sigma c_i^2}{(\alpha_i+\sigma t)^3}\right)^2}{\left(\sum_{i=1}^{n}\frac{c_i^2}{(\alpha_i+\sigma t)^2}\right)^2} + \frac{2}{p-2}\frac{1}{t^2}.$$

Define two vectors in $\mathbb{R}^n$ by

$$a = \left(\frac{\sqrt{6}\sigma c_1}{(\alpha_1+\sigma t)^2}, \ldots, \frac{\sqrt{6}\sigma c_n}{(\alpha_n+\sigma t)^2}\right)^{\mathrm{T}}, \quad b = \left(\frac{c_1}{\alpha_1+\sigma t}, \ldots, \frac{c_n}{\alpha_n+\sigma t}\right)^{\mathrm{T}}.$$

Applying Cauchy–Schwartz inequality, we obtain

$$\begin{aligned}\left(\sum_{i=1}^{n}\frac{2\sigma c_i^2}{(\alpha_i+\sigma t)^3}\right)^2 &< (a^{\mathrm{T}}b)^2\\ &\leq (a^{\mathrm{T}}a)(b^{\mathrm{T}}b)\\ &= \left(\sum_{i=1}^{n}\frac{6\sigma^2 c_i^2}{(\alpha_i+\sigma t)^4}\right)\left(\sum_{i=1}^{n}\frac{c_i^2}{(\alpha_i+\sigma t)^2}\right).\end{aligned}$$

Therefore, $p''(t) > 0$ for all t such that $p(t)$ is well-defined. It follows that $p(t)$ is strictly convex for $t \in (\max\{-\alpha_2/\sigma, 0\}, -\alpha_1/\sigma)$. Thus, $p(t)$, as well as $h(t)$, has at most two real roots in this interval. Let $t_1 < t_2$ be the only two roots of $h(t)$. Suppose $h'(t_1) > 0$ and $h'(t_2) > 0$. Then, for sufficiently small $\epsilon \in (0, (t_2 - t_1)/2)$, we have

$$h(t_1 + \epsilon) > h(t_1) = 0, \; h(t_2 - \epsilon) < h(t_2) = 0.$$

Therefore, there is a $\tilde{t} \in [t_1 + \epsilon, t_2 - \epsilon]$ such that $h(\tilde{t}) = 0$, which is a contradiction. Consequently, the secular function $h(t)$ has at most one real root satisfying $h'(t) > 0$. Following Theorem 3.2, the proof is complete. ∎

Remark 2 It is indeed surprising that we can obtain the necessary and sufficient conditions for the local non-global minimizers for general $p > 2$, and the number of which is at most one. It was previously known to be true for $p = 4$ in the double well potential function [21] but the technique to generalize the result is non-trivial as we do not have a convex secular function for $2 < p < 4$.

4. Application to (p-RS) subject to linear inequality constraints for $p = 4$

As an application of Theorems 3.2 and 3.3, the model (p-RS$_m$) specified in (6)-(7) for the special case $p = 4$ is studied in this section. We first show that two combinatorial optimization problems, one of which is the partition problem and the other is the binary quadratic problem, can be reduced to special cases of (p-RS$_m$) with $p = 4$. Secondly, we shall show that (p-RS$_m$) is in general NP-hard, but for $p = 4$ and m a fixed constant, it can be solved in polynomial time.

The partition problem (PP) is to ask whether the following equation

$$r^{\mathrm{T}}x = 0, \; x \in \{-1, 1\}^n \tag{30}$$

has a solution, where $r = (r_1, \ldots, r_n)^{\mathrm{T}}$ is a given vector of positive integer entries. We show that the partition problem (PP) can be equivalently formulated as the following (p-RS$_m$) with $m = n$:

$$\begin{aligned} (p\text{-RS}_n) \;\; \min \quad & \phi(x) = x^{\mathrm{T}}\left(\frac{1}{nr^{\mathrm{T}}r}rr^{\mathrm{T}} - \frac{4}{n}I\right)x + \frac{4}{p(p-1)n^{p/2}}\|x\|^p \\ \text{s.t.} \quad & x \in [-1, 1]^n. \end{aligned} \tag{31}$$

To see this, notice that

$$\frac{1}{r^{\mathrm{T}}r}rr^{\mathrm{T}} \preceq I$$

and

$$\|x\| \le \sqrt{n}, \quad xx^{\mathrm{T}} \preceq nI, \; \forall x \in [-1, 1]^n,$$

It follows that $\phi(x)$ is concave since

$$\begin{aligned} \nabla^2\phi(x) &= \frac{2}{nr^{\mathrm{T}}r}rr^{\mathrm{T}} - \frac{8}{n}I + \frac{4}{p(p-1)n^{p/2}}(p\|x\|^{p-2}I + p(p-2)\|x\|^{p-4}xx^{\mathrm{T}}) \\ &\preceq \frac{2}{n}I - \frac{8}{n}I + \frac{4}{p(p-1)n^{p/2}}(pn^{p/2-1}I + p(p-2)n^{p/2-2}nI) \\ &= -\frac{2}{n}I \prec 0. \end{aligned}$$

Then, $v(p\text{-RS}_n)$ must be attained at one of the vertices of $[-1, 1]^n$. More precisely, we have

$$v(p\text{-RS}_n) = \frac{1}{nr^{\mathrm{T}}r}\min_{x\in\{-1,1\}^n}(r^{\mathrm{T}}x)^2 - 4 + \frac{4}{p(p-1)}.$$

Consequently, $v(p\text{-RS}_n) = -4 + 4/p(p-1)$ if and only if (PP) as in (30) has a solution. Since (PP) is NP-hard [9], it implies that (p-RS$_m$) with $m = n$ is NP-hard.

In general, we have

THEOREM 4.1 *For any* $p > 2$, (p-RS$_n$), $\bigcup_{m\geq n}$ ($p-$RS$_m$), $\bigcup_{m\in\mathbb{N}}$ ($p-$RS$_m$) *are all NP-hard.*

As a second example to reformulate a general combinatorial optimization problem to (p-RS$_m$), we consider the following binary quadratic optimization:

$$\text{(BQP)}\ d^* = \min \quad x^{\mathrm{T}} Qx \tag{32}$$

$$\text{s.t.} \quad e^{\mathrm{T}} x = k, \quad Ax \leq b, \quad x \in \{0,1\}^n, \tag{33}$$

where Q is a symmetric $n \times n$ matrix, k is an integer, $A \in \mathbb{R}^{m\times n}$ and $b \in \mathbb{R}^m$ have integer entries. Let a_j^{T} be the jth row of A for $j = 1, \ldots, m$. Without loss of generality, we assume that the relaxed region $S = \{x \in [0,1]^n : e^{\mathrm{T}} x = k,\ Ax \leq b\}$ is non-empty and $\|a_j\|_1 \geq 2$ for $j = 1, \ldots, m$, see also [15]. To avoid triviality, we also assume that $n \geq 2$. Define $A_{\max} = \max_{1\leq j\leq m} \|a_j\|_\infty$ and

$$G_A = \begin{cases} \dfrac{A_{\max}^2}{A_{\max} - 1}, & \text{if } A_{\max} \geq 2, \\ 2, & \text{if } A_{\max} = 1. \end{cases}$$

As a direct corollary of Lemma 2 in [15], we have

LEMMA 4.2 [15] *Let x be a vertex of S and $x \notin \{0,1\}^n$. Then*

$$x^{\mathrm{T}}(e - x) \geq \frac{1}{G_A}.$$

To see (BQP) can be reformulated as ($p-$RS$_m$), we first define, for any given $\theta_1 \geq 0,\ \theta_2 \geq 0$,

$$d(x) = x^{\mathrm{T}} Qx + \theta_1(e^{\mathrm{T}} x - x^{\mathrm{T}} x) + \theta_2(e^{\mathrm{T}} x - x^{\mathrm{T}} x)^2,$$

and observe that

$$d^* = \min_{e^{\mathrm{T}} x=k,\ Ax\leq b,\ x\in\{0,1\}^n} x^{\mathrm{T}} Qx = \min_{e^{\mathrm{T}} x=k,\ Ax\leq b,\ x\in\{0,1\}^n} \mathrm{d}(x).$$

Secondly, define the continuous relaxation of (BQP) as

$$d^c = \min_{e^{\mathrm{T}} x=k,\ Ax\leq b,\ x\in[0,1]^n} x^{\mathrm{T}} Qx \leq d^*.$$

Notice that, if $\theta_1 > \max\{0,\ n\theta_2 + \lambda_{\max}(Q)\},\ \theta_2 \geq 0$, then it holds that, for any $x \in [0,1]^n$,

$$\begin{aligned} \nabla^2 d(x) &= 2Q - 2\theta_1 I + 2\theta_2(e - 2x)(e - 2x)^{\mathrm{T}} - 4\theta_2(e^{\mathrm{T}} x - x^{\mathrm{T}} x)I \\ &\preceq 2Q - 2\theta_1 I + 2n\theta_2 I \prec 0, \end{aligned}$$

which implies that $d(x)$ with $\theta_1 > \max\{0,\ n\theta_2 + \lambda_{\max}(Q)\},\ \theta_2 \geq 0$ is a strictly concave function over $x \in [0,1]^n$.

Let V be the set of all vertices of $S = \{x \in [0,1]^n : e^{\mathrm{T}}x = k,\ Ax \le b\}$. By choosing

$$\theta_2 \ge G_A^2(d^* - d^c) \ge 0 \quad \text{and} \quad \theta_1 > \max\{0, n\theta_2 + \lambda_{\max}(Q)\},$$

the function $d(x)$ is concave and $\theta_1(e^{\mathrm{T}}x - x^{\mathrm{T}}x) \ge 0$ on $[0,1]^n$. Then,

$$\begin{aligned}
&\min_{e^{\mathrm{T}}x=k,\ Ax\le b,\ x\in[0,1]^n} \mathrm{d}(x)\\
&= \min_{x\in V} \mathrm{d}(x)\\
&= \min\left\{\min_{V\setminus\{0,1\}^n} \mathrm{d}(x),\ \min_{V\cap\{0,1\}^n} \mathrm{d}(x)\right\}\\
&\ge \min\left\{\min_{V\setminus\{0,1\}^n} \{x^{\mathrm{T}}Qx + \theta_1\left(e^{\mathrm{T}}x - x^{\mathrm{T}}x\right)\} + \min_{V\setminus\{0,1\}^n} \theta_2\left(e^{\mathrm{T}}x - x^{\mathrm{T}}x\right)^2,\ d^*\right\}\\
&\ge \min\left\{\min_{e^{\mathrm{T}}x=k,\ Ax\le b,\ x\in[0,1]^n} x^{\mathrm{T}}Qx + \min_{V\setminus\{0,1\}^n} \theta_2(e^{\mathrm{T}}x - x^{\mathrm{T}}x)^2,\ d^*\right\}\\
&\ge \min\left\{d^c + \frac{\theta_2}{G_A^2},\ d^*\right\}\\
&= d^*\\
&\ge \min_{e^{\mathrm{T}}x=k,\ Ax\le b,\ x\in[0,1]^n} \mathrm{d}(x),
\end{aligned} \tag{34}$$

where (34) follows from Lemma 4.2. Therefore, we have reduced (BQP) to the following special case of (RS_{m+n+1}) with $p=4$:

$$\begin{aligned}
\min \quad & d(x) = x^{\mathrm{T}}(Q - \theta_1 I - 2\theta_2 kI)x + \theta_2\|x\|^4 + \theta_1 k + \theta_2 k^2\\
\text{s.t.} \quad & e^{\mathrm{T}}x = k, \quad Ax \le b, \quad x \in [0,1]^n.
\end{aligned}$$

Finally, we remark that (BQP) covers some applications as its special case. The first one is k-dispersion-sum problem:

$$\begin{aligned}
\text{(KDSP)} \quad & \max x^{\mathrm{T}}Dx\\
\text{s.t.} \quad & e^{\mathrm{T}}x = k, \quad x \in \{0,1\}^n.
\end{aligned}$$

The problem allocates k facilities at part of n predefined locations ($k \le n$) in a way that the distance sum among the k established facilities is maximized, where the distance sum is specified by a given distance matrix D and $x^{\mathrm{T}}Dx$. It is known that the k-dispersion-sum problem (KDSP) is NP-hard, even if the distance matrix satisfies the triangle inequality. See, for example, [7,12]. The second application of (BQP) is the quadratic assignment problem, one of the great challenges in combinatorial optimization:

$$\begin{aligned}
\text{(QAP)} \quad & \min\ \mathrm{trace}(FXDX^{\mathrm{T}})\\
\text{s.t.} \quad & Xe = X^{\mathrm{T}}e = e, \quad X \in \{0,1\}^{n\times n},
\end{aligned}$$

where trace$(\cdot)$ is the trace of the matrix $(\cdot)$, F and D correspond to the flow and distance matrices, respectively. Notice that (QAP) can be reformulated as a special case of (BQP):

$$\begin{aligned}
\text{(QAP)}\ \min \quad & \mathrm{trace}(FXDX^{\mathrm{T}})\\
\text{s.t.} \quad & e^{\mathrm{T}}Xe = n, \quad Xe \le e, \quad X^{\mathrm{T}}e \le e, \quad X \in \{0,1\}^{n\times n}.
\end{aligned}$$

Since (QAP) is NP-hard [19], so is (BQP).

4.1 *Polynomially solvable cases*

We show that, for any fixed positive integer m, ($p-$RS$_m$) with $p=4$ is in Class P. The reduction argument used to prove the result is inherited from [14].

Let $L=\{x \mid \; l_i \le a_i^{\mathrm{T}} x \le u_i, \; i=1,\dots,m\}$ and X_0^* be the set of the global minimizers of (p-RS). Since X_0^* is either a singleton or a sphere, we can first check whether $X_0^* \bigcap L$ is empty in polynomial time by the following lemma:

LEMMA 4.3 [14] *Let $A \in R^{m\times q}$ and $b \in R^m$, where m is fixed and q is arbitrary. For any given $r>0$, it is polynomially checkable whether $\{u \in R^q \mid Au \le b, \|u\|^2 = r\}$ is empty. Moreover, if the set is non-empty , a feasible point can be found in polynomial time.*

If $X_0^* \bigcap L \neq \emptyset$, any point in the intersection is a global minimizer of ($p-$RS$_m$).

If this is not the case, the global minimizer of ($p-$RS$_m$), denoted by x_m^*, would reside either in the interior of L satisfying $l_i < a_i^{\mathrm{T}} x_m^* < u_i$, or on the boundaries satisfying one of the equalities: $a_i^{\mathrm{T}} x_m^* = l_i$, $a_i^{\mathrm{T}} x_m^* = u_i$ for $i=1,2,\dots,n$. In the former case, x_m^* must be the unique local non-global minimizer $\underline{x}_0$ of (p-RS).

Define, for the interior case, that

$$\delta^0 = \begin{cases} g(\underline{x}_0), & \text{if } l_i < {a_i}^{\mathrm{T}} \underline{x}_0 < u_i, \; \forall i, \\ \infty, & \text{otherwise}, \end{cases}$$

while for the boundary cases

$$\begin{aligned} \delta_m^{j_1} := \min \quad & g(x) = \frac{1}{2} x^{\mathrm{T}} H x + c^{\mathrm{T}} x + \frac{\sigma}{4}\|x\|^4 \\ \text{s.t.} \quad & {a_j}^{\mathrm{T}} x = l_j, \\ & l_i \le {a_i}^{\mathrm{T}} x \le u_i, \; i = 1,\dots,j-1, j+1,\dots,m, \end{aligned} \tag{35}$$

$$\begin{aligned} \delta_m^{j_2} := \min \quad & g(x) = \frac{1}{2} x^{\mathrm{T}} H x + c^{\mathrm{T}} x + \frac{\sigma}{4}\|x\|^4 \\ \text{s.t.} \quad & {a_j}^{\mathrm{T}} x = u_j, \\ & l_i \le {a_i}^{\mathrm{T}} x \le u_i, \quad i = 1,\dots,j-1, j+1,\dots,m. \end{aligned} \tag{36}$$

It follows immediately that

$$g(x_m^*) = \min\{\delta^0, \delta_m^{1_1}, \delta_m^{1_2}, \dots, \delta_m^{m_1}, \delta_m^{m_2}\}.$$

Now we observe that, for each $\delta_m^{j_1}$ (similarly for $\delta_m^{j_2}$), the equality constraint $a_j^{\mathrm{T}} x = l_j$ (or $a_j^{\mathrm{T}} x = u_j$) can be used to eliminate one variable and, most importantly due to the speciality of $p=4$, the p-regularized structure with $p=4$ is retained. As a result, we come to a $n-1$ dimensional (p-RS$_{m-1}$).

Let $P_j \in \mathbb{R}^{n\times(n-1)}$ be a column-orthogonal matrix such that $a_j^{\mathrm{T}} P_j = 0$ and z_0 a feasible solution to (35). Then $z_0 - P_j P_j^{\mathrm{T}} z_0$ is also feasible to (35). Using the null-space representation, we have

$$\{x \in \mathbb{R}^n \mid {a_j}^{\mathrm{T}} x = b_j\} = \{z_0 - P_j {P_j}^{\mathrm{T}} z_0 + P_j z \mid z \in \mathbb{R}^{n-1}\}.$$

Then,

$$\begin{aligned}\|x\|^4 &= ((z_0 - P_jP_j{}^{\mathrm{T}}z_0 + P_jz)^{\mathrm{T}}(z_0 - P_jP_j{}^{\mathrm{T}}z_0 + P_jz))^2\\ &= (z_0{}^{\mathrm{T}}(I - P_jP_j{}^{\mathrm{T}})(I - P_jP_j{}^{\mathrm{T}})z_0 + 2z_0{}^{\mathrm{T}}(I - P_jP_j{}^{\mathrm{T}})P_jz + z^{\mathrm{T}}P_j^{\mathrm{T}}P_jz)^2\\ &= (z_0{}^{\mathrm{T}}(I - P_jP_j{}^{\mathrm{T}})z_0 + z^{\mathrm{T}}z)^2\\ &= (z_0{}^{\mathrm{T}}(I - P_jP_j{}^{\mathrm{T}})z_0)^2 + 2(z_0{}^{\mathrm{T}}(I - P_jP_j{}^{\mathrm{T}})z_0)z^{\mathrm{T}}z + \|z\|^4,\end{aligned}$$

and (35) is reduced to a form of ($p-\mathrm{RS}_{m-1}$):

$$\begin{aligned}\min \quad & g(z_0 - P_jP_j{}^{\mathrm{T}}z_0 + P_jz)\\ \text{s.t.} \quad & l_i \le a_i{}^{\mathrm{T}}(z_0 - P_jP_j{}^{\mathrm{T}}z_0 + P_jz) \le u_i, \quad i = 1, \ldots, m,\ i \ne j.\end{aligned}$$

We can repeat the reduction scheme until there is no more linear inequality constraint or $n = 1$ (when $n < 2m$). In the former case, it is an unconstrained p-RS, while in the latter case the polyhedron reduces to an interval and there can be at most four critical points (two boundary points plus two local minimizers at best). The total number of reductions depends on m exponentially, but for a fixed m, it is a constant factor. We thus arrive the following conclusion:

THEOREM 4.4 *For each fixed m,* ($p-\mathrm{RS}_m$) *with* $p = 4$ *is polynomially solvable.*

5. Conclusions

Our comprehensive analysis on the p-RSs for general $p > 2$ gives the most detailed comparison between (TRS) and (p-RS); and between (ETRS) and ($p-\mathrm{RS}_m$). We virtually confirm that the weighted regularization term $(\sigma/p)\|x\|^p$ can be freely exchanged with the trust region constraint $\|x\|^2 \le \Delta$ in almost all applications. The interchange can have meaningful implications. Many approximate models such as (SQP) which previously go with the trust region method can now be considered to incorporate with the p-RS. In some cases when the trust region constraint $\|x\|^2 \le \Delta$ introduces enormous difficulty, for example, the CDT problem [4], our study suggests that one may consider (p-RS) subject to just one ellipsoid constraint, instead of the intersection of one ball and one ellipsoid. More future researches are of course needed.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This research was supported by National Natural Science Foundation of China under grants 11331012, 11321061, 11471325 and 11571029, by Taiwan MOST 103-2115-M-006-014-MY2, and by fundamental research funds for the Central Universities under grant YWF-16-BJ-Y-11.

References

[1] A. Beck and A. Ben-Tal, *On the solution of the Tikhonov regularization of the total least squares problem*, SIAM J. Optim. 17(1) (2006), pp. 98–118.
[2] D. Bienstock and A. Michalka, *Polynomial solvability of variants of the trust-region subproblem*, ACM-SIAM Symposium on Discrete Algorithms, 2014, pp. 380–390.

[3] C. Cartis, N.I.M. Gould, and Ph.L. Toint, *Adaptive cubic regularisation methods for unconstrained optimization. Part I: Motivation, convergence and numerical results*, Math. Program. 127 (2011), pp. 245–295.
[4] M.R. Celis, J.E. Dennis, and R.A. Tapia, *A trust region algorithm for nonlinear equality constrained optimization*, in *Numerical Optimization*, P.T. Boggs, R.H. Byrd, and R.B. Schnabel, eds., SIAM, Philadelphia, 1985, pp. 71–82.
[5] A.R. Conn, N.I.M. Gould, and Ph.L. Toint, *Trust-Region Methods*, MPS-SIAM Series on Optimization Vol. 01, Philadelphia, SIAM, 2000.
[6] W. Dinkelbach, *On nonlinear fractional programming*, Manage. Sci. 13 (1967), pp. 492–498.
[7] E. Erkut, *The discrete p-dispersion problem*, European J. Oper. Res. 46 (1990), pp. 48–60.
[8] S.C. Fang, D. Gao, G.X. Lin, R.L. Sheu, and W. Xing, *Double well potential function and its optimization in the n-dimensional real space – Part I*, J. Ind. Manag. Optim. (2016). doi:10.3934/jimo.2016073.
[9] M.R. Garey and D.S. Johnson, *Computers and Intractability: A Guide to the Theory of NP-Completeness*, Freeman, San Francisco, 1979.
[10] N.I.M. Gould, D.P. Robinson, and H. Sue Thorne, *On solving trust-region and other regularised subproblems in optimization*, Math. Program. Comput. 2 (2010), pp. 21–57.
[11] A. Griewank, *The modification of Newton's method for unconstrained optimization by bounding cubic Terms*, Technical Report DAMTP/NA12, Department of Applied Mathematics and Theoretical Physics, Cambridge University, Cambridge, UK, 1981.
[12] P. Hansen and I. Moon, *Dispersing Facilities on a Network*, Presentation at the TIMS/ORSA Joint National Meeting, Washington, DC, 1988.
[13] R.A. Horn and C.R. Johnson, *Matrix Analysis*, Cambridge University Press, Cambridge, 1985.
[14] Y. Hsia and R.L. Sheu, *Trust region subproblem with a fixed number of additional linear inequality constraints has polynomial complexity*, preprint (2013). arXiv:1312.1398.
[15] Y. Hsia and Y. Wang, *A new penalty parameter for linearly constrained 0-1 quadratic programming problems*, Optim. Lett. 7(4) (2013), pp. 765–778.
[16] R.L. Jerrard, *Lower bounds for generalized Ginzburg–Landau functionals*, SIAM J. Math. Anal. 30(4) (1999), pp. 721–746.
[17] J.M. Martínez, *Local minimizers of quadratic function on Euclidean balls and spheres*, SIAM J. Optim. 4 (1994), pp. 159–176.
[18] Y. Nesterov and B.T. Polyak, *Cubic regularization of Newton method and its global performance*, Math. Program. 108 (2006), pp. 177–205.
[19] S. Sahni and T. Gonzalez, *P-complete approximation problems*, J. Assoc. Comput. Mach. 23 (1976), pp. 555–565.
[20] M. Weiser, P. Deuflhard, and B. Erdmann, *Affine conjugate adaptive Newton methods for nonlinear elastomechanics*, Optim. Methods Softw. 22 (2007), pp. 413–431.
[21] Y. Xia, R.L. Sheu, S.C. Fang, and W. Xing, *Double well potential function and its optimization in the n-dimensional real space – Part II*, J. Ind. Manag. Optim. (2016). doi:10.3934/jimo.2016074.
[22] Y.X. Yuan, *Recent advances in trust region algorithms*, Math. Program. Ser. B 151 (2015), pp. 249–281.