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

Ya-xiang Yuan

State Key Laboratory of Scientific and Engineering Computing
Inst of Comput. Math. and Scientific/Engineering Computing
Academy of Mathematics and Systems Science
Chinese Academy of Sciences, Beijing 100190, China

joint work with *Yong Xia, Ruey-lin Sheu*
Optim. Methods Softw. 32 (2017), no. 5, 1059-1077.

Outline

Outline

Outline

Outline

Outline

Introduction

$$\min_{x \in \Re^n} c^T x + \frac{1}{2} x^T H x + \frac{\sigma}{p} \|x\|_2^p \qquad (p - RS)$$

$\sigma > 0$, $p > 2$. $H \in \Re^{n \times n}$ symmetric.

Subproblem for Unconstrained Optimization: $\min f(x)$

$$f(x) \approx f(x_k) + f'(x_k)(x - x_k) + \frac{1}{2}(x - x_k)^T H_k (x - x_k)$$

Regularization term: (Penalty term)

$$||x - x_k||_2^p$$

Why $p > 2$?

Self Adaptive Property of $||x||^p$

Define $g(x) = c^T x + \frac{1}{x} x^T H x + \frac{\sigma}{p} ||x||^p$. As a penalty function, $g(x)$

- When $||x|| \to +\infty$,

$$\lim_{||x|| \to +\infty} g(x) = +\infty.$$

- When $x \to 0$,

$$g(x) \sim c^T x + \frac{1}{2} x^T H x$$

Previous Works

Cubic Regularization: $p = 3$

$$\min_{x \in \Re} c^T x + \frac{1}{2} x^T H x + \frac{\sigma}{3} \|x\|_2^3$$

Griewank (1981)
Nesterov and Polyak (2006)
Weiser, Deuflhard and Erdmann (2007)
Cartis, Gould and Toint (2011)

quartic regularization: $p = 4$,
Fang, Gao, Lin Sheu, Xing (2012)

general $p-$ **regularization** for $p > 2$:
Gould, Robinson and Thorne (2010)

Related Problem: Trust Region Subproblem

Trust region Subproblem

$$\text{(TRS)} \quad \min \quad \frac{1}{2}x^T Hx + c^T x$$
$$\text{s.t.} \quad \|x\|_2^2 \le \Delta, \; x \in \mathbb{R}^n.$$

Quadratic regularization: (Levenberg-Marquardt Technique)

$$\min \frac{1}{2}x^T Hx + c^T x + \frac{\sigma}{2}||x||_2^2 \qquad (2 - RS)$$

Relations between (TRS) and (2-RS): Gay(1981), Moré and Sorenson(1983).

Interesting Properties of (TRS) Local minimizers, global minimizers Martinez(1994): necessary conditions for local non-global minimizers.

Motivation

Properties of (p-RS) for general $p > 2$:

- Conditions for local minimizers and global minimizers
- conditions for local non-global minimizers
- number of local non-global minimizers

Notations.

- Diag(x): diagonal matrix with components $x_1, \ldots, x_n$.
- $\lambda_i(P)$ is the ith smallest eigenvalue of P
- $\|x\| = \|x\|_2$

Necessary Condition for a Local Minimizer

Lemma

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

$$\nabla g(\underline{x}) = \left(H + \sigma\|\underline{x}\|^{p-2} I\right) \underline{x} + c = 0, \quad (1)$$

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

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

direct calculations

Necessary and Sufficient Conditions for Global Minimizer

Theorem

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

Necessary: See Gould, Robinson and Thorne(2010).

N.I.M. Gould, D.P. Robinson and H.S. Thorne, On solving trust region and other regularised subproblems in optimization, *Mathematical Programming Computation*, 2(2010) pp.21-57.

Sufficient:

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, $x^T H x \geq 0, \ \forall x \in \mathbb{R}^n$. Thus, $x^* = 0_n$ is a global minimizer since

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

Now, Assume $\|x^*\| > 0$.

Define $Q = H + \sigma\|x^*\|^{p-2} I$. $Q \succeq 0$. For any $x \in \mathbb{R}^n$ and $x \neq x^*$,

$$\begin{aligned} g(x) &= \frac{1}{2}x^T H x + c^T x + \frac{\sigma}{p}\|x\|^p \\ &= \frac{1}{2}x^T Q x + c^T x + \frac{\sigma}{p}\|x^*\|^p \left(\left(\frac{\|x\|^2}{\|x^*\|^2} \right)^{\frac{p}{2}} - \frac{p}{2}\frac{\|x\|^2}{\|x^*\|^2} \right) \end{aligned} \quad (3)$$

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

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

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,

$$g(x) \quad \geq \quad \frac{1}{2}x^T Q x + c^T x + \frac{\sigma}{p}\|x^*\|^p(1 - \frac{p}{2}). \tag{4}$$

By $Q \succeq 0$, the lower bounding function of g in the right hand side of (4) 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 (4). As a consequence,

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

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

Finally, from (3), if $\hat{x}$ is also a global minimizer of (p-RS), $\hat{x}$ must minimize both $\frac{1}{2}x^T Qx + c^T x$ and $\left(\frac{\|x\|^2}{\|x^*\|^2}\right)^{\frac{p}{2}} - \frac{p}{2}\frac{\|x\|^2}{\|x^*\|^2}$ simultaneously. This can happen if and only if $Q\hat{x} = -c$ and $\|\hat{x}\| = \|x^*\|$.

Set of Global Minimizers

Theorem

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

Ideas of Proof

Consider Diagonal H

Assume H is diagonal, i.e.,

$$H = \mathrm{Diag}(\alpha_1, \ldots, \alpha_n), \tag{5}$$

where

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

and k is the multiplicity of the smallest eigenvalue α_1.
Otherwise let $H = U \Sigma U^T$, and let $y = U^T x$

Outlines of the Proof

- Define $t^* = \|x^*\|^{p-2} \geq 0$, independent of x^*.
- $t^* \in (\max\{-\frac{\alpha_1}{\sigma}, 0\}, +\infty)$.
- Solution x^* satisfying $(H + \sigma t^* I)x^* = -c$

If $H + \sigma t^* I$ is invertible

- $x_i^* = \frac{-c_i}{\alpha_i + \sigma t^*}, \ i = 1, \ldots, n.$
- $$h(t) = \sum_{i=1}^{n} \frac{c_i^2}{(\alpha_i + \sigma t)^2} - t^{\frac{2}{p-2}}, \ t \in I_g = \left(\max\{-\frac{\alpha_1}{\sigma}, 0\}, +\infty\right). \quad (6)$$

  $h(t)$ can only have a unique root on I_g, which must be t^*.

If $H + \sigma t^* I$ is singular

- $t^* = \frac{-\alpha_1}{\sigma}. \Longrightarrow (\alpha_1 \le 0)$
- $c_1^2 + \ldots + c_k^2 = 0$, and $\alpha_i + \sigma t^* > 0, \ i = k+1, k+2, \ldots, n$ such that

  $$\hat{x}^* = \left(0, 0, \ldots, 0, \frac{-c_{k+1}}{\alpha_{k+1} - \alpha_1}, \ldots, \frac{-c_n}{\alpha_n - \alpha_1}\right)^T \quad (7)$$

  is one trivial solution to $(H - \alpha_1 I)x^* = -c$.
- $\hat{h}(t) = \sum_{i=k+1}^{n} \frac{c_i^2}{(\alpha_i + \sigma t)^2} - t^{\frac{2}{p-2}}, \ t \in I_{\hat{g}} = \left[-\frac{\alpha_1}{\sigma}, +\infty\right).$

  $\hat{h}(t)$: strictly decreasing on $I_{\hat{g}}$ and $\lim_{t \to +\infty} \hat{h}(t) = -\infty$.

If $h\left(-\frac{\alpha_1}{\sigma}\right) = 0$

- $t^* = \frac{-\alpha_1}{\sigma}$ is the only root of $\hat{h}(t)$ on $I_{\hat{g}}$.
- $\hat{x}^*$ defined in (7) is the unique global minimizer of (p-RS).

If $\hat{h}\left(-\frac{\alpha_1}{\sigma}\right) < 0$

- $\hat{h}(t) = 0$ has no solution
- the trivial solution $\hat{x}^*$ to $(H - \alpha_1 I)x^* = -c$ does not satisfy $t^* = \frac{-\alpha_1}{\sigma} = \|\hat{x}^*\|^{p-2}$.
- Any x^* satisfying

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

is a global minimizer of (p-RS).

Hidden Convexity

Lemma

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

$$c_i x_i^* \leq 0, \; i = 1, \ldots, n.$$

Proof

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

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

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

By the above lemma,

$$\begin{array}{ll} \min & \sum_{i=1}^{n}\{\frac{\alpha_i}{2}x_i^2 + c_i x_i\} + \frac{\sigma}{p}\left(\sum_{i=1}^{n} x_i^2\right)^{\frac{p}{2}} \\ \text{s.t.} & c_i x_i \le 0, \quad i = 1, \ldots, n. \end{array} \tag{8}$$

and (p-RS) share the same optimal solution set.
Using the nonlinear one-to-one map:

$$x_i = \begin{cases} \sqrt{z_i}, & \text{if } c_i \le 0, \\ -\sqrt{z_i}, & \text{if } c_i > 0, \end{cases} \quad i = 1, \ldots, n, \tag{9}$$

we convert the problem (8) into the following convex program:

$$\begin{array}{ll} \min & -\sum_{i=1}^{n} |c_i|\sqrt{z_i} + \frac{1}{2}\sum_{i=1}^{n}\alpha_i z_i + \frac{\sigma}{p}\left(\sum_{i=1}^{n} z_i\right)^{\frac{p}{2}} \\ \text{s.t.} & z_i \ge 0, \ i = 1, \ldots, n. \end{array} \tag{10}$$

Ramark

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

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

- For $p \geq 4$, $h(t)$ is strictly decreasing on I_g and convex.
- For $p = 3$, if $H \not\succeq 0$ (which ensures that $c \neq 0$) and h is restricted to a finite subinterval of I_g covering t^*, by properly choosing the regularization parameter σ, $h(t)$ can be made convex.
- Notice that the secular function for (TRS) is always convex.

Necessary Condition for Local Non-global Minimizer

Assume

- H is diagonal.
- $\alpha_1 < 1$
- $n \geq 2$

Lemma

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

$$\alpha_2 + \sigma \|\underline{x}\|^{p-2} > 0. \tag{11}$$

Proof of the Lemma

$\underline{x}$ is a local but non-global minimizer, $H + \sigma\|\underline{x}\|^{p-2} \not\succeq 0$. Thus, $\alpha_1 + \sigma\|\underline{x}\|^{p-2} < 0$.

Second order necessary conditions imply

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

This gives $x_1 \neq 0$ and

$$(\alpha_1 + \sigma\|\underline{x}\|^{p-2})(\underline{x}_2)^2 + (\alpha_2 + \sigma\|\underline{x}\|^{p-2})(\underline{x}_1)^2 \geq 0.$$

Therefore $\alpha_2 + \sigma\|\underline{x}\|^{p-2} \geq 0$.

But $\alpha_1 + \sigma\|\underline{x}\|^{p-2} < 0$, thus, $\alpha_1 \neq \alpha_2$.

Prove $\alpha_2 + \sigma\|\underline{x}\|^{p-2} > 0$ **by contradiction.** Assume $\alpha_2 + \sigma\|\underline{x}\|^{p-2} = 0$.

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

By $\alpha_1 - \alpha_2 < 0$ and

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

it implies that $\underline{x}_2 = 0$ and, first order necessary condition gives

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

w.l.o.g. assume $c_1 > 0$ and $\underline{x}_1 > 0$. Define

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

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

$$\gamma(t) = \{(k(t), t, \underline{x}_3, \ldots, \underline{x}_n) \mid t \in [-\underline{x}_1, \underline{x}_1]\}. \tag{13}$$

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

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

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

$$\frac{d^i}{dt^i} g(\gamma(0)) = 0 (i = 1, 2, 3), \ \frac{d^4}{dt^4} g(\gamma(0)) = -\frac{3(\alpha_2 - \alpha_1)}{\underline{x}_1^2} < 0.$$

$\square$

Necessary and Sufficient Conditions for local non-global Minimizers

Theorem

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

$$\underline{x} = -\left(H + \sigma \underline{t}^* I\right)^{-1} c, \tag{14}$$

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

$$h(t) = \left\| \left(H + \sigma t I\right)^{-1} c \right\|^2 - t^{\frac{2}{p-2}}, \; t \in \left(\max\left\{-\frac{\alpha_2}{\sigma}, 0\right\}, -\frac{\alpha_1}{\sigma}\right) \tag{15}$$

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

Proof: Necessary

local non-global minimizer $\underline{x}$ satisfies

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

First order necessary condition gives

$$\underline{x}_i = \frac{-c_i}{\alpha_i + \sigma\|\underline{x}\|^{p-2}}, \ i = 1, \ldots, n \tag{16}$$

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

Define $\underline{t}^* = \|\underline{x}\|^{p-2}$. $h(\underline{t}^*) = 0$ and

$$t^* \in \left(\max\left\{-\frac{\alpha_2}{\sigma}, 0\right\}, -\frac{\alpha_1}{\sigma}\right).$$

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

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

necessary optimality condition is equivalent to

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

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,

$$\begin{aligned} 0 &\leq \det\left(\mathrm{Diag}(-1, 1, \ldots, 1) + \sigma(p-2)\|\underline{x}\|^{p-4}(\Gamma \underline{x})(\Gamma \underline{x})^T\right) \\ &= (\frac{p}{2} - 1)\|\underline{x}\|^{p-4} h'(\|\underline{x}\|^{p-2}) = (\frac{p}{2} - 1)\|\underline{x}\|^{p-4} h'(\underline{t}^*). \end{aligned}$$

Thus, $h'(t^*) \geq 0$

Prove $h'(t^*) > 0$ by contradiction.

If $h'(t^*) = 0$, There exist $u = (u_1, \ldots, u_n)^T \neq 0$ such that

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

Define $q(\beta) := g(\underline{x} + \beta u)$. We can verify that

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

$q'(0) = q''(0) = 0$. and $(q'''(0))^2 > 0$.

Contradiction! Necessary part is proved.

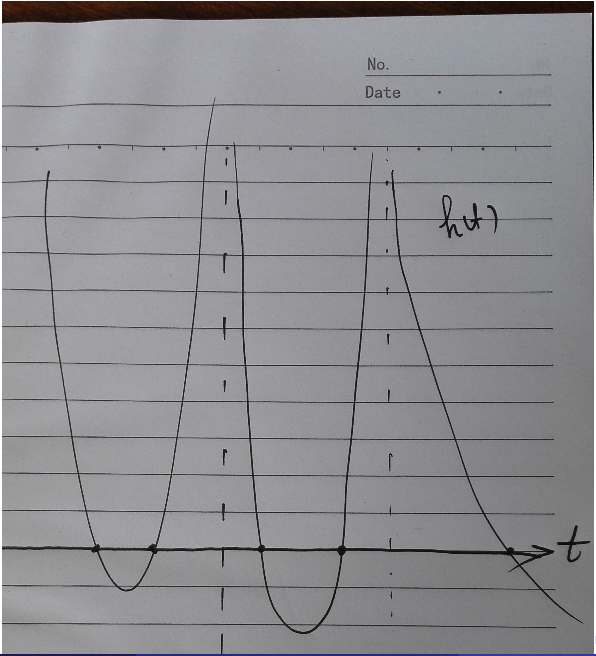
Sufficient Part

Main Idea

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

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

$$\nabla^2 g(\underline{x})>0$$


No.
Date
h(t)
t

Only 1 Local Non-Global Minimizer

Theorem

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

Idea of Proof

- Define the function

$$p(t) = \log\left(\| (H + \sigma t I)^{-1} c\|^2\right) - \frac{2}{p-2}\log(t),$$

$$t \in \left(\max\left\{-\frac{\alpha_2}{\sigma}, 0\right\}, -\frac{\alpha_1}{\sigma}\right).$$

- $p(t) = 0$ and $h(t) = 0$ equivalent
- $p(t)$ convex

Remark

It is indeed surprising that we obtain

- *the necessary and sufficient conditions for the local non-global minimizers for general* $p > 2$,
- *the number of local non-global minimizer is at most one.*

It was previously known to be true for $p = 4$ *in the double well potential function but the technique to generalize the result is non-trivial as we do not have a convex secular function for* $2 < p < 4$.

Application to Linear Inequality Constraints

linearly constrained p-regularized problem:

$$
\begin{array}{lll}
(\mathrm{p}-\mathrm{RS}_m) & \min & \frac{1}{2}x^T Hx + c^T x + \frac{\sigma}{p}\|x\|^p \\
& \text{s.t.} & l_i \le a_i^T x \le u_i, \ i = 1, \ldots, m,
\end{array}
$$

where $l_i \le u_i \in \mathbb{R}$ for $i = 1, \ldots, m$.

Partition Problem

Example: Given numbers:

1, 2, 3, 6, 8.

Partition them into

1, 3, 6

2, 8

Partition Problem

Given $r = (r_1, r_2, ..., r_n) \in \Re^n_+$, *Solve*

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

Relation between Partition Problem and $(p - RS_n)$

Partition Problem can be formulated into:

Special $(p - RS_n)$ problem:

$$\min \phi(x) = x^T \left(\frac{1}{nr^T r} rr^T - \frac{4}{n} I \right) x + \frac{4}{p(p-1)n^{p/2}} ||x||^p$$

$$s.t. \quad -1 \leq x_i \leq 1, \quad i = 1, 2, ..., n.$$

because $\phi(x)$ is concave in $[-1, 1]^n$

NP hardness of $(\mathrm{p} - \mathrm{RS}_m)$

Partition Problem is NP-hard

Theorem

For any $p > 2$,

$$(\mathrm{p} - \mathrm{RS}_n),$$

$$\bigcup_{m \geq n} (\mathrm{p} - \mathrm{RS}_m),$$

$$\bigcup_{m \in \mathcal{N}} (\mathrm{p} - \mathrm{RS}_m)$$

are all NP-hard.

Binary Quadratic Optimization

$$\begin{array}{ll} \min & x^T Qx \\ \text{s. t.} & e^T x = k,\ Ax \leq b,\ x \in \{0,1\}^n \end{array}$$

where $Q \in \Re^{n\times n}$ is a symmetric matrix, k is an integer, $A \in \Re^{m\times n}$ and $b \in \Re^m$ have integer entries.

reduced to a special $(\mathrm{PS_{m+n+1}})$ with p=4:

$$\min d(x) = x^T(Q - \theta_1 I - 2\theta_2 k\mathrm{I})\mathrm{x} + \theta_2||\mathrm{x}||^4 + \theta_1\mathrm{k} + \theta_2\mathrm{k}^2$$

$$s.t. \quad e^T x = k, \quad Ax \leq b, \quad x \in [0,\ 1]^n$$

Special Cases of Binary Quadratic Optimization

$k-$**dispersion-sum problem**

$$\max x^T D x$$
$$s.t. \quad e^T x = k, \quad x \in \{0,1\}^n .$$

Quadratic Assignment Problem

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

both **KDSP** and **QAP** are NP-hard.

Polynomial Solvable for Problems with fixed m

Theorem

For each fixed m, (p-RS_m) *with* $p = 4$ *is polynomially solvable.*

Idea of Proof

- Reducing $(\mathrm{p} - \mathrm{RS}_k)$ to $(\mathrm{p} - \mathrm{RS}_{k-1})$
- based on the following result:

Lemma (Xia and Sheu)

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* $\{u \in R^q \mid Au \leq b, \|u\|^2 = r\}$ *is empty. Moreover, if the set is nonempty, a feasible point can be found in polynomial time.*

Conclusions

What we have done:

1. complete analysis to the *p*-regularized subproblems for general $p > 2$ gives the most detailed comparison between the p-regularized subproblems and trust region subproblems, including
   - necessary and sufficient conditions for minimizers
   - conditions for local non-global minimizers
   - number of local non-global minimizers
2. For $p = 4$, (p-RS) can be extended to solve the *p*-regularized subproblems subject to additional linear constraints.

Thanks very much!