Jordan-algebraic aspects of nonconvex optimization over symmetric cones

L.Faybusovich* and Ye LU†

Abstract. We illustrate the usefulness of Jordan algebraic technique for nonconvex optimization by considering a potential-reduction algorithm for a nonconvex quadratic function over the domain obtained as the intersection of a symmetric cone with an affine subspace.

Key words. Jordan algebras, symmetric cone, nonconvex optimization.

1 Introduction

Jordan-algebraic technique proved to be very useful for the analysis of convex optimization problems over symmetric cones. See e.g. [f1,f2,f3,f4,fT M,S,T]. Since the class of symmetric cones contains the positive orthant in R^n, the second-order cone and cone of positive-definite symmetric matrices, this technique is applicable to a broad class of optimization problems. However, it is also quite useful for nonconvex optimization problems over symmetric cones. To illustrate this point we have chosen a potential-reduction algorithm for the minimization of a nonconvex quadratic function developed and analyzed by Y.Ye[Ye1] for the case of polyhedral constraints. Many other algorithms (e.g. path-following algorithms based on trust-region ideas) can be analyzed using the same technique. An excellent introduction to the theory of Euclidean Jordan algebras suitable for our purposes is contained in [FK]. See also [f1]. Alternatively, the reader may keep in mind the following concrete

*Department of Mathematics, University of Notre Dame. 255 Hurley Hall, Notre Dame, IN 46556, USA. email: lfaybuso@nd.edu

†Department of Mathematics, University of Notre Dame. 255 Hurley Hall, Notre Dame, IN 46556, USA. email: ylu4@nd.edu

situation. The Jordan algebra V is the vector space of symmetric matrices with the multiplication.

$$X \circ Y = \frac{XY+YX}{2}, \tag{$*$}$$

where in the right-hand side of $(*)$ XY is the standard matrix multiplication. Then $P(X)Y = XYX$, and in this example the spectral decomposition in Jordan-algebraic sense coincides with the usual spectral decomposition (the Jordan-algebraic determinant is the usual matrix determinant as well).

2 Symmetric cones and Jordan algebras

In this section we briefly describe some relevant Jordan-algebraic concepts. For details see [FK].

Let V be an Euclidean Jordan algebra and Ω be a cone of invertible squares in V. We define $\langle x, y \rangle = tr(x \circ y)$ as the canonical scalar product.

Let $F(x) = -\log\det(x)$, $x \in \Omega$. Then $F'(x) = \bigtriangledown F(x) = -x^{-1}$, $F''(x) = H_F(x) = P(x)^{-1}$, here $H_F(x)$ is the Hession of F evaluated at $x \in \Omega$ with respect to the canonical scalar product $\langle , \rangle$; x^{-1} is the inverse of x and $P(x)$ is the quadratic representation of x.

Given $x \in \Omega$, we define a local norm $\|v\|_x = \langle v, P(x)^{-1}v\rangle^{\frac{1}{2}}$, for $v \in V$. Let $B_x(y, r)$ denote the open ball of radius r centered at y, where the radius is measured w.r.t. $\| \; \|_x$. Let $P(x)^{-\frac{1}{2}}v = \sum_{i=1}^{r} \lambda_i e_i$, $\lambda_1 \le \lambda_2 \le \cdots \le \lambda_r$ be the spectral decomposition. Then $\|P(x)^{-\frac{1}{2}}v\|^2 = \sum_{i=1}^{r} \lambda_i^2$, and we define $\|P(x)^{-\frac{1}{2}}v\|_\infty := \max_{1\le i\le r}\{|\lambda_i|\}$. The following lemma is well-known. For a Jordan-algebraic proof see e.g.[f1].

Lemma 2.1. *Given $x \in \Omega$, we have $B_x(x, 1) \subseteq \Omega$.*

Lemma 2.2. *Given $x \in \Omega$ and $\|P(x)^{-\frac{1}{2}}\xi\| < 1$, then for $F(x) = -\ln\det(x)$, we have*

$$|F(x+\xi)-F(x)-\langle F'(x),\xi\rangle-\frac{\langle\xi,F''(x)\xi\rangle}{2}|\leq\frac{\|P(x)^{-\frac{1}{2}}\xi\|_\infty\|\xi\|_x^2}{3(1-\|P(x)^{-\frac{1}{2}}\xi\|_\infty)}.$$

Proof. We have $F'(x)=-x^{-1}$, $F''(x)=P(x)^{-1}$ and

$$\begin{aligned}\Delta: &= F(x+\xi)-F(x)-\langle F'(x),\xi\rangle-\frac{\langle\xi,F''(x)\xi\rangle}{2}\\ &= \int_0^1\int_0^t\langle\xi,(F''(x+s\xi)-F''(x))\xi\rangle dsdt\\ &= \int_0^1\int_0^t\langle P(x)^{-\frac{1}{2}}\xi,(P(x)^{\frac{1}{2}}(P(x+s\xi)^{-1}-P(x)^{-1})P(x)^{\frac{1}{2}})P(x)^{-\frac{1}{2}}\xi\rangle dsdt\\ &= \int_0^1\int_0^t\langle P(x)^{-\frac{1}{2}}\xi,(P(e+sP(x)^{-\frac{1}{2}}\xi)^{-1}-I)P(x)^{-\frac{1}{2}}\xi\rangle dsdt,\end{aligned}$$

here in the last equality we use the properties $P(x)^{-\frac{1}{2}}x=e$ and $P(P(y)x)=P(y)P(x)P(y)$ whose proof can be found in [Faraut].
Let $P(x)^{-\frac{1}{2}}\xi=\sum_{i=1}^{i}\lambda_ie_i$, $\lambda_1\leq\lambda_2\leq\cdots\leq\lambda_r$ be the spectral decomposition. Then we know that $\sum_{i=1}^{r}\lambda_i^2<1$ since we have assumed $\|P(x)^{-\frac{1}{2}}\xi\|<1$. The eigenvalues of $P(e+sP(x)^{-\frac{1}{2}}\xi)^{-1}$ have the form $\frac{1}{(1+s\lambda_i)(1+s\lambda_j)}$, $i,j=1,2,\cdots r$ (see e.g.[f2]). Therefore we have $(\frac{1}{(1+s\lambda_r)^2}-1)I\preceq P(e+sP(x)^{-\frac{1}{2}}\xi)^{-1}-I\preceq(\frac{1}{(1+s\lambda_1)^2}-1)I$, Thus

$$\begin{aligned}\|\xi\|_x^2\int_0^1\int_0^t(\frac{1}{(1+s\lambda_r)^2}-1)dsdt &\leq \Delta\leq\|\xi\|_x^2\int_0^1\int_0^t(\frac{1}{(1+s\lambda_1)^2}-1)dsdt\Rightarrow\\ \|\xi\|_x^2\int_0^1-\frac{t^2\lambda_r}{1+t\lambda_r}dt &\leq \Delta\leq\|\xi\|_x^2\int_0^1-\frac{t^2\lambda_1}{1+t\lambda_1}dt\Rightarrow\\ \|\xi\|_x^2\int_0^1-\frac{t^2\|P(x)^{-\frac{1}{2}}\xi\|_\infty}{1-t\|P(x)^{-\frac{1}{2}}\xi\|_\infty}dt &\leq \Delta\leq\|\xi\|_x^2\int_0^1\frac{t^2\|P(x)^{-\frac{1}{2}}\xi\|_\infty}{1-t\|P(x)^{-\frac{1}{2}}\xi\|_\infty}dt\Rightarrow\\ |\Delta| &\leq \|P(x)^{-\frac{1}{2}}\xi\|_\infty\|\xi\|_x^2\int_0^1\frac{t^2}{1-t\|P(x)^{-\frac{1}{2}}\xi\|_\infty}dt\\ &\leq \frac{\|P(x)^{-\frac{1}{2}}\xi\|_\infty\|\xi\|_x^2}{3(1-\|P(x)^{-\frac{1}{2}}\xi\|_\infty)}.\end{aligned}$$

Remark. Since it is easy to see that $\|P(x)^{-\frac{1}{2}}\xi\|_\infty\leq\|\xi\|_x$, We have $|F(x+\xi)-F(x)-\langle F'(x),\xi\rangle-\frac{\langle\xi,F''(x)\xi\rangle}{2}|\leq\frac{\|\xi\|_x^3}{3(1-\|\xi\|_x)}$. This result can be found in [Renegar].

We need the following. For proof see e.g. [FK].

Lemma 2.3. *(a)*$P(x)x^{-1} = x$*;*
(b)$\langle x^{-1}, x\rangle = r$*, here* $r = rank(V)$*.*

Lemma 2.4. $P(x)$ *is a linear automorphism of* Ω *for each* $x \in \Omega$*, that is* $P(x)\Omega = \Omega$*.*

3 Optimality conditions

The detailed analysis of optimality conditions for problems with conic constraints can be found in [BS]. Here we present a direct proof of first-order optimality conditions sufficient for our purposes.

Consider the following optimization problem:

$$\begin{aligned} (P1)\min \quad & f(x) \\ \text{subject to} \quad & x \in a + X, \\ & x \in K. \end{aligned}$$

Here $f : E \to R$ is a differentiable function, $a \in E$, X is a vector subspace in E and K is a cone in E. We assume that $K \cap (a + X)$ is bounded and $K^{\circ} \cap (a + X)$ is nonempty; here $K^{\circ} = int(K)$, is the interior of K. Now we are ready to formulate our first-order optimal condition.

Theorem 3.1. *(First-order optimality condition). If* x^* *is a local minimum solution of (P1), then* $\exists s \in K^*$ *such that* $f'(x^*) - s \in X^{\perp}$ *and* $\langle x^*, s\rangle = 0$*. Here* K^* *is the cone dual to* K*.*

Lemma 3.1. *Given* $x \in K$*, let* $K(x) = \{y \in E | x + ty \in K$*, for sufficiently small* $t > 0\}$*, then its dual* $K(x)^* = \{s \in K^* | \langle s, x\rangle = 0\}$*.*

Proof. One can easily see that $K(x)$ is a convex cone. Since $K \subset K(x)$, we have $K(x)^* \subset K^*$. Let $s \in K^*, \langle s, x\rangle = 0$. If $y \in K(x)$, $\exists t > 0$ such that $x^* + ty \in K$. Then, $\langle s, x + ty\rangle = t\langle s, y\rangle \geq 0$. Hence, $\langle s, y\rangle \geq 0$. Thus $\{s \in K^* | \langle s, x\rangle = 0\} \subset K(x)^*$.
Inversely, let $s \in K(x)^* \subset K^*$, We have $\langle s, x\rangle \geq 0$. But $-x \in K(x)$, Indeed, if $0 < t < 1$, $x + t(-x) = (1 - t)x \in K$, i.e. $-x \in K(x)$. But then $\langle s, -x\rangle \geq 0$, i.e. $\langle s, x\rangle \leq 0$. Therefore, $\langle s, x\rangle = 0$. We conclude $K(x)^* \subset \{s \in K^* | \langle s, x\rangle = 0\}$, which completes the proof. □

Proof of Theorem 3.1. Let $K(x^*) := \{y \in E | x^* + ty \in K$, for sufficiently small $t > 0\}$ and $\pi : E \to X^\perp$ be the orthogonal projector. Let $N = \{x \in R \times X^\perp : x = (t, y), t = \langle f'(x^*), p\rangle, y = \pi(p), p \in K(x^*)\}$ and $P = \{(t, y) \in R \times X^\perp : t < 0, y = 0\}$.
We claim that $N \cap P = \emptyset$. Indeed, if $N \cap P \neq \emptyset$, $\exists p \in K(x^*) \cap X$ such that $\langle f'(x^*), p\rangle < 0$. But then $x^* + tp \in (a + X) \cap K$ for sufficiently small $t > 0$. Consider $\varphi(t) = f(x^* + tp)$, we have $\varphi'(0) = \langle f'(x^*), p\rangle < 0$. Hence $\varphi(t) < \varphi(0) = f(x^*)$ for sufficiently small $t > 0$. It contradicts to the fact that x^* is a local minimum. Thus $N \cap P = \emptyset$. By separation theorem $\exists (r, u) \in R \times X^\perp, (r, u) \neq 0$, such that

$$r\langle f'(x^*), p\rangle + \langle u, \pi(p)\rangle \geq rt \tag{3.1}$$

for any $p \in K(x^*)$ and $t < 0$. In particular, $\langle rf'(x^*) + u, p\rangle \geq 0, \forall p \in K(x^*)$, i.e. $rf'(x^*) + u \in K(x^*)^*$ for some $u \in X^\perp$. This can be rewritten in the form $rf'(x^*) - s \in X^\perp$ for some $s \in K(x^*)^*$. Observe that the inequality (3.1) implies $r \geq 0$ (If $r < 0$, consider $t \to -\infty$). From Lemma 3.1, we know that $s \in K^*$ and $\langle x^*, s\rangle = 0$.
Now it remains to show the case $r = 0$ is impossible. Indeed, if $r = 0$, then $\exists s \in K^* \cap X^\perp$ such that $\langle s, x^*\rangle = 0$. On the other hand, by our assumptions, $\exists y \in (a + X) \cap K^\circ$. But $x^* = y + (x^* - y)$ and $x^* - y \in X$. Hence, $\langle s, y\rangle = \langle s, x^*\rangle = 0$ (since $s \in X^\perp$). But since $y \in K^\circ$, $\langle s, y\rangle = 0$ implies $s = 0$, which is impossible. Thus $r > 0$, which completes the proof. □

In the remaining part of this paper, we always consider the following optimization problem.

$$
\begin{aligned}
(P)\min \quad & q(x) = \frac{\langle x, Qx\rangle}{2} + \langle c, x\rangle \\
\text{subject to} \quad & x \in a + X, \\
& x \in \overline{\Omega}.
\end{aligned}
$$

Here $Q : V \mapsto V$ is a symmetric linear operator on an Euclidean Jordan algebra V, X is a subspace of V and $c \in V$, $\overline{\Omega} = \{x^2 | x \in V\}$ and Ω is the cone of invertible squares in V; it is known that $\Omega = int(\overline{\Omega})$. We assume Q is not positive semi-definite on the subspace X.

Let $\underline{z}$ and $\overline{z}$ be the minimal and maximal values of the objective function $q(x)$ over the feasible set. We define x as an ϵ-KKT solution for problem (P) if $\exists s \in \overline{\Omega}$ such that $Qx + c - s \in X^{\perp}$ and $\frac{\langle x,s\rangle}{\overline{z}-\underline{z}} \leq \epsilon$.

Similarly, we say that x is an ϵ-minimizer of problem (P) if $\exists s \in \Omega$, such that $Qx + c - s \in X^{\perp}$ and $\frac{q(x)-\underline{z}}{\overline{z}-\underline{z}} \leq \epsilon$.
We note that the minimizer of problem (P) is a special KKT point such that $q(x) = \underline{z}$.

4 Potential reduction algorithm.

In this section we present a potential- reduction algorithm for problem (P). The complexity analysis of the algorithm is presented. The total number of iterations of the algorithm will be a polynomial in the rank of V.

Consider, first, the ball-constrained quadratic programming problem:

$$
\begin{aligned}
(BP)\min \quad & q(x) = \frac{\langle x, Qx\rangle}{2} + \langle c, x\rangle \\
\text{subject to} \quad & x \in X, \\
& \|x\|^2 \leq \alpha^2.
\end{aligned}
$$

Here, $0 < \alpha < 1$ is a real parameter. The following necessary and sufficient optimality conditions are well known (see e.g. [Ye1]). There exists real

μ such that

$$\begin{aligned}(Q + \mu I)x + c &= z \; for \; some \; z \in X^{\perp}, \\ x &\in X, \\ \mu &\geq \max\{0, -\lambda\}, \\ \|x\| &= \alpha.\end{aligned}$$

Here $\lambda = \underline{\lambda}(\pi Q \pi)$ denote the least eigenvalue of $\pi Q \pi$, where π is an orthogonal projector of V onto the subspace X. Since we have assumed that Q is not positive semi-definite on the subspace X, we must have $\lambda < 0$.

Assuming that $z \leq \underline{z}$ is a lower bound of $q(x)$ on the feasible set of the problem(P), we define the following potential function

$$\Psi(x) = \rho \ln(q(x) - z) + F(x),$$

here $F(x) = -\ln\det(x)$ and $\rho > 0$ is a constant parameter. We assume $\mathcal{F}_p = \overline{\Omega} \cap (a + X)$ is bounded and $\Omega \cap (a + X)$ is nonempty; x_0 is the analytic center of $\mathcal{F}_p$, that is the minimal value point of $F(x)$ over $\mathcal{F}_p$. Without loss of generality, we can assume $x_0 = e$, here e is the identity element of Jordan algebra. The algorithm starts from x_0, and will generate a sequence of $x_k \in \mathcal{F}_p$ such that $\Psi(x_{k+1}) < \Psi(x_k)$. We have

$\rho \ln(q(x_k) - z) - \rho \ln(q(x_0) - z) \leq \rho \ln(q(x_k) - z) + F(x_k) - \rho \ln(q(x_0) - z) - F(x^0) = \Psi(x_k) - \Psi(x_0)$.

Here the first inequality holds because we start from the analytic center of $\mathcal{F}_p$ so that $F(x_k) \geq F(x_0)$. Suppose that

$$\Psi(x_k) - \Psi(x_0) \leq \rho \ln \epsilon \tag{4.1}$$

Then we have

$$\frac{q(x_k) - \underline{z}}{\overline{z} - \underline{z}} \leq \frac{q(x_k) - \underline{z}}{q(x_0) - \underline{z}} \leq \frac{q(x_k) - z}{q(x_0) - z} \leq \epsilon.$$

We will try to construct an algorithm with the goal to decrease the value of $\Psi(x)$. We denote $\Delta_k = q(x_k) - z$, and select $d_{x_k} \in X$ such that $x^{k+1} = x_k + d_{x_k} \in \Omega$. Then,

$\rho\ln(q(x_k+d_{x_k})-z)-\rho\ln(q(x_k)-z)=\rho\ln(\Delta_k+\frac{\langle d_{x_k},Qd_{x_k}\rangle}{2}+\langle Qx_k+c,d_{x_k}\rangle)-\rho\ln\Delta_k=\rho\ln(1+\frac{\frac{\langle d_{x_k},Qd_{x_k}\rangle}{2}+\langle Qx_k+c,d_{x_k}\rangle}{\Delta_k})\leq\frac{\rho}{\Delta_k}(\frac{\langle d_{x_k},Qd_{x_k}\rangle}{2}+\langle Qx_k+c,d_{x_k}\rangle)$.

From Lemma 2.1. we know that if $\|P(x)^{-\frac{1}{2}}d_{x_k}\|\leq\alpha<1$, $x^{k+1}=x_k+d_{x_k}\in\Omega$ and from Lemma 2.2 we have

$$\begin{aligned}F(x_{k+1})-F(x_k) &\leq \langle -x^{-1},d_{x_k}\rangle+\frac{\langle d_{x_k},P(x_k)^{-1}d_{x_k}\rangle}{2}+\frac{\|d_{x_k}\|_{x_k}^3}{3(1-\|d_{x_k}\|_{x_k})}\\ &\leq \langle -x^{-1},d_{x_k}\rangle+\frac{\alpha^2}{2}+\frac{\alpha^3}{3(1-\alpha)}\leq\langle -x^{-1},d_{x_k}\rangle+\frac{\alpha^2}{2(1-\alpha)}.\end{aligned}$$

Therefore, we get

$$\Psi(x_{k+1})-\Psi(x_k)\leq\frac{\rho}{\Delta_k}(\frac{\langle d_{x_k},Qd_{x_k}\rangle}{2}+\langle Qx_k+c-\frac{\Delta_k}{\rho}x^{-1},d_{x_k}\rangle)+\frac{\alpha^2}{2(1-\alpha)}.$$

We choose d_{x^k} as a solution to the following problem.

$$\begin{aligned}\min\quad & \frac{\langle d_{x_k},Qd_{x_k}\rangle}{2}+\langle Qx_k+c-\frac{\Delta_k}{\rho}(x_k)^{-1},d_{x_k}\rangle\\ s.t.\quad & d_{x_k}\in X,\\ & \|P(x_k)^{-\frac{1}{2}}d_{x_k}\|^2\leq\alpha^2.\end{aligned}$$

Let $d'_{x_k}=P(x_k)^{-\frac{1}{2}}d_{x_k}$, $Q_k=P(x_k)^{\frac{1}{2}}QP(x_k)^{\frac{1}{2}}$, $c_k=P(x_k)^{\frac{1}{2}}Qx_k+P(x^k)^{\frac{1}{2}}c-\frac{\Delta_k}{\rho}P(x_k)^{\frac{1}{2}}x_k^{-1}$. We can transform the above problem to standard ball-constrained problem as the following form

$$\begin{aligned}(BQP)\min\quad & q'(d'_{x_k})=\frac{\langle d'_{x_k},Q_kd'_{x_k}\rangle}{2}+\langle c_k,d'_{x_k}\rangle\\ s.t.\quad & d'_{x_k}\in P(x_k)^{-\frac{1}{2}}X,\\ & \|d'_{x_k}\|^2\leq\alpha^2.\end{aligned}$$

Using optimality conditions formulated above, we obtain that there exists real μ_k such that

$$\begin{aligned}(Q_k+\mu^kI)d'_{x_k}+c^k &= z_k,\ here\ z_k\in(P(x_k)^{-\frac{1}{2}}X)^\perp=P(x_k)^{\frac{1}{2}}X^\perp, \quad (4.2)\\ d'_{x_k} &\in P(x_k)^{-\frac{1}{2}}X,\\ \mu_k &\geq \max\{0,-\lambda_k\},\\ \|d'_{x_k}\| &= \alpha.\end{aligned}$$

Here, λ_k is the least eigenvalue of $\pi_k Q_k \pi_k$, where π_k is an orthogonal projector from V onto $P(x_k)^{-\frac{1}{2}}X$. Since we have assumed that Q is not positive semi-definite on the subspace X and $P(x_k)^{-1}$ is positive definite for every x_k, we must have that Q_k is not positive semi-definite on the subspace $P(x_k)^{-\frac{1}{2}}X$. Therefore, $\lambda_k < 0$.

Let $s(\mu_k) = Q(x_k + d_{x_k}) + c - P(x_k)^{-\frac{1}{2}}z_k$, $p_k = Q_k d'_{x_k} + c_k - z_k = P(x_k)^{\frac{1}{2}}Qd_{x_k} + P(x_k)^{\frac{1}{2}}Qx_k + P(x_k)^{\frac{1}{2}}c - \frac{\Delta_k}{\rho}P(x_k)^{\frac{1}{2}}(x_k)^{-1} - P(x_k)^{\frac{1}{2}}P(x_k)^{-\frac{1}{2}}z_k = P(x_k)^{\frac{1}{2}}s(\mu_k) - \frac{\Delta_k}{\rho}P(x_k)^{\frac{1}{2}}x_k^{-1}$. Then from (4.2) we have $\mu_k = \frac{\|p_k\|}{\alpha}$ and $d'_{x_k} = -\frac{\alpha p_k}{\|p_k\|}$. Also,

$$\begin{aligned} q'(d'_{x_k}) &= \frac{\langle d'_{x_k}, Q_k d'_{x_k}\rangle}{2} + \langle c_k, d'_{x_k}\rangle = \langle d'_{x_k}, Q_k d'_{x_k} + c_k\rangle - \frac{\langle d'_{x_k}, Q_k d'_{x_k}\rangle}{2} \\ &= \langle d'_{x_k}, Q_k d'_{x_k} + c_k - z_k\rangle - \frac{\langle d'_{x_k}, Q_k d'_{x_k}\rangle}{2} = -\alpha\|p_k\| - \frac{\langle d'_{x_k}, Q_k d'_{x_k}\rangle}{2} \\ &= -\alpha^2\mu_k - \frac{\langle d'_{x_k}, Q_k d'_{x_k}\rangle}{2} \le -\alpha_2\mu_k - \frac{\lambda_k}{2}\|d'_{x_k}\|^2 \le -\frac{\alpha^2\mu_k}{2} = -\frac{\alpha\|p_k\|}{2}. \end{aligned}$$

This implies that

$$\Psi(x_{k+1}) - \Psi(x_k) \le \frac{\rho}{\Delta_k}q'(d'_{x_k}) + \frac{\alpha^2}{2(1-\alpha)} \le -\frac{\alpha\rho}{2\Delta_k}\|p_k\| + \frac{\alpha^2}{2(1-\alpha)}.$$

Here we can see that if $\frac{\rho}{\Delta_k}\|p_k\| \ge \frac{3}{4}$ and α is chosen around $\frac{1}{4}$, then

$$\Psi(x_{k+1}) - \Psi(x_k) \le -\frac{5}{96}.$$

Using (4.1), we obtain:

Theorem 4.1. *Let $\alpha = \frac{1}{4}$ and p_k, Δ_k be defined as above . Then, If $\frac{\rho}{\Delta_k}\|p_k\| \ge \frac{3}{4}$ for all k, the algorithm returns an ϵ-minimizer in at most $\frac{96\rho\ln\frac{1}{\epsilon}}{5}$ iterations.*

Let us consider the case, where $\frac{\rho}{\Delta_k}\|p_k\| < \frac{3}{4}$ at some k. We are going to show that by choosing a suitable ρ and z we can guarantee that the next iteration x_{k+1} of the algorithm is an ϵ-KKT.

Actually we can address a weaker case, where $\frac{\rho}{\Delta_k}\|p_k\| < 1$, that is,

$\|P(x_k)^{\frac{1}{2}}\frac{\rho}{\Delta_k}s(\mu_k)-P(x_k)^{\frac{1}{2}}x_k^{-1})\| = \|P(x_k)^{-\frac{1}{2}}(P(x_k)\frac{\rho}{\Delta_k}s(\mu_k)-P(x_k)x_k^{-1})\| = \|P(x_k)^{-\frac{1}{2}}(P(x_k)\frac{\rho}{\Delta_k}s(\mu_k)-x_k)\| < 1.$

The last equality follows from Lemma 2.3 part (a). Then from Lemma 2.1, we know that $P(x_k)\frac{\rho}{\Delta_k}s(\mu_k) \in \Omega$, Lemma 2.4 implies $\frac{\rho}{\Delta_k}s(\mu_k) \in \Omega$, and consequently $s(\mu_k) \in \Omega$.

Define $t = P(x_k)^{\frac{1}{2}}\frac{\rho}{\Delta_k}s(\mu_k) - (P(x_k)^{\frac{1}{2}}x_k^{-1})$, then $s(\mu_k) = \frac{\Delta_k}{\rho}(P(x_k)^{-\frac{1}{2}}t + x_k^{-1})$. and

$$
\begin{aligned}
\langle x_k, s(\mu_k)\rangle &= \frac{\Delta_k}{\rho}(\langle P(x_k)^{-\frac{1}{2}}x_k, t\rangle + \langle x_k, x_k^{-1}\rangle)\\
&\leq \frac{\Delta_k}{\rho}(\|P(x_k)^{-\frac{1}{2}}x_k\|\|t\| + \langle x_k, x_k^{-1}\rangle)\\
&\leq \frac{\Delta_k}{\rho}(\langle x_k, P(x_k)^{-1}x_k\rangle^{\frac{1}{2}} + \langle x_k, x_k^{-1}\rangle)\\
&= \frac{\Delta_k}{\rho}(\langle x_k, x_k^{-1}\rangle^{\frac{1}{2}} + \langle x_k, x_k^{-1}\rangle) = \frac{\Delta_k}{\rho}(\sqrt{r} + r).
\end{aligned}
$$

The last two equalities follow from Lemma 2.3. Similarly, we have

$$
\begin{aligned}
\langle d_{x_k}, s(\mu_k)\rangle &= \langle P(x_k)^{\frac{1}{2}}d'_{x_k}, s(\mu_k)\rangle = \frac{\Delta_k}{\rho}(\langle d'_{x_k}, t\rangle + \langle d'_{x_k}, P(x_k)^{\frac{1}{2}}x_k^{-1}\rangle)\\
&\leq \frac{\Delta_k}{\rho}(\|d'_{x_k}\|\|t\| + \|d'_{x_k}\|\|P(x_k)^{\frac{1}{2}}x_k^{-1}\|)\\
&\leq \frac{\Delta_k}{\rho}(\alpha + \alpha\langle x_k^{-1}, P(x_k)x_k^{-1}\rangle^{\frac{1}{2}})\\
&= \frac{\Delta_k}{\rho}(\alpha + \alpha\langle x_k^{-1}, x_k\rangle^{\frac{1}{2}}) = \frac{\Delta_k}{\rho}(\alpha + \alpha\sqrt{r}).
\end{aligned}
$$

Therefore, combining the above two inequalities, we have

$$
\langle x_{k+1}, s(\mu_k)\rangle = \langle x_k, s(\mu_k)\rangle + \langle d_{x_k}, s(\mu_k)\rangle \leq \frac{\Delta_k}{\rho}(r + (1+\alpha)\sqrt{r} + \alpha). \quad (4.3)
$$

Then we can see that

$\frac{\langle x_{k+1}, s(\mu_k)\rangle}{\Delta_k} \leq \frac{r+(1+\alpha)\sqrt{r}+\alpha}{\rho}$.

If we know $\underline{z}$, we can choose $\rho = \frac{r+(1+\alpha)\sqrt{r}+\alpha}{\epsilon}$ to make sure that $\frac{\rho}{\Delta_k}\|p_k\| < 1$ implies x_{k+1} is an ϵ-KKT. But usually, we don't know the exact $\underline{z}$. We overcome this problem by choosing a suitable lower bound z. Assuming that x_0 is the analytic center of Ω, the following lemma tells us that Ω is contained between two ellipsoids centered at x_0.

Lemma 4.1. *Let $F(x) = -\ln det(x)$ and x_0 is the analytic center of $\overline{\Omega} \cap (a+X)$, then*

$$B_{x_0}(x_0,1) \cap (a+X) \subseteq \overline{\Omega} \cap (a+X) \subseteq B_{x_0}(x_0,r).$$

Proof. The leftmost inclusion follows directly from Lemma 2.1. Now we need to prove the rightmost inclusion. Given $y \in \overline{\Omega} \cap (a+X)$, consider the spectral decomposition $P(x_0)^{-\frac{1}{2}}y = \sum_{i=1}^{i} \lambda_i e_i$. Since x_0 is a minimizer of $F(x) = -\ln det(x)$ over $\overline{\Omega} \cap (a+X)$, $F'(x_0) = -x^{-1} \in X^{\perp}$. Therefore $0 = \langle x_0^{-1}, x_0 - y\rangle = \langle P(x_0)^{-\frac{1}{2}}e, x_0 - y\rangle = \langle e, e - P(x_0)^{-\frac{1}{2}}y\rangle = r - tr(P(x_0)^{-\frac{1}{2}}y) = r - \sum_{i=1}^{r} \lambda_i$.
On the other hand, $\langle y - x_0, P(x_0)^{-1}(y-x_0)\rangle = \langle P(x_0)^{-\frac{1}{2}}y - e, P(x_0)^{-\frac{1}{2}}y - e\rangle = \sum_{i=1}^{r}(\lambda_i - 1)^2$. To get an upper bound of the radius of the ball, we consider the optimization problem: $\max \sum_{i=1}^{r}(\lambda_i - 1)^2$, subject to $\sum_{i=1}^{r}\lambda_i = r$ and $\lambda_i \geq 0$, $i = 1, \cdots r$. We know this is a concave optimization problem and the optimal solution is one of the extreme points which are of the form $(0, \cdots, 0, r, 0, \cdots, 0)$. Therefore, $\max \sum_{i=1}^{r}(\lambda_i - 1)^2 = (r-1)^2 + r - 1 = r(r-1) < r^2$. This completes the proof. ☐

We need the following lemma which is due to Y.Ye [Ye2].

Lemma 4.2. *Given $r' > 0$, let $d(r')$ be the minimizer for*

$$\begin{aligned} \min \quad & q(d) = \frac{\langle d, Qd\rangle}{2} + \langle c, d\rangle \\ s.t. \quad & d \in X, \\ & \|d\|^2 \leq r'^2. \end{aligned}$$

Then, for $0 < r' \leq R$, $q(0) - q(d(r')) \geq \frac{r'^2}{R^2}(q(0) - q(d(R)))$.

Consider the following ball-constrained problem:

$$\begin{aligned} \min \ q(x) - q(x_0) \quad & = \frac{\langle x - x_0, Q(x-x_0)\rangle}{2} + \langle Qx_0 + c, x - x_0\rangle \\ s.t. \quad & x - x_0 \in X, \\ & \|x - x_0\|_{x_0} \leq r'. \end{aligned}$$

Since we have assumed that $x_0 = e$ and we know $P(e) = I$, we can apply Lemma 4.2 directly. We let $\widehat{x}(r')$ be the minimizer of the above problem determined by r', then $q(x_0) - q(\widehat{x}(1)) \geq \frac{1}{r^2}(q(x_0) - q(\widehat{x}(r')) \geq \frac{1}{r^2}(q(x_0) - \underline{z})$, here the first inequality follows from Lemma 4.2 and second inequality is from Lemma 4.1. Thus we can set a lower bound as

$$z = q(x_0) - r^2(q(x_0) - q(\widehat{x}(1))).$$

Then, $\Delta_k = q(x_k) - q(x_0) + r^2(q(x_0) - q(\widehat{x}(1))) \leq (1 + r^2)(\overline{z} - \underline{z})$.

Therefore we have

$$\frac{\langle x_{k+1}, s(\mu_k)\rangle}{\overline{z} - z} \leq \frac{(1+r^2)\langle x_{k+1}, s(\mu_k)\rangle}{\Delta_k} \leq \frac{(1+r^2)(r+(1+\alpha)\sqrt{r}+\alpha)}{\rho},$$

the last inequality follows from (4.3). Hence

Theorem 4.2. *If we know $\underline{z}$, we set $z = \underline{z}$ and choose $\rho = \frac{r+(1+\alpha)\sqrt{r}+\alpha}{\epsilon}$; If we don't know $\underline{z}$, we set $z = q(x_0) - r^2(q(x_0) - q(\widehat{x}(1)))$ and choose $\rho = \frac{(1+r^2)(r+(1+\alpha)\sqrt{r}+\alpha)}{\epsilon}$. For both of these two cases, The condition $\frac{\rho}{\Delta_k}\|p_k\| < 1$ implies x_{k+1} is an ϵ-KKT.*

We can now summarize the potential-reduction algorithm:

Algorithm 4.1. If we know $\underline{z}$, we set $z = \underline{z}$ and choose $\rho = \frac{r+(1+\alpha)\sqrt{r}+\alpha}{\epsilon}$; If we don't know $\underline{z}$, we set $z = q(x_0) - r^2(q(x_0) - q(\widehat{x}(1)))$ and choose $\rho = \frac{(1+r^2)(r+(1+\alpha)\sqrt{r}+\alpha)}{\epsilon}$, set $\alpha = \frac{1}{4}$ and $k = 0$.
If $\frac{\rho}{\Delta_k}\alpha\mu_k = \frac{\rho}{\Delta_k}\|p_k\| < \frac{3}{4}$ or $\frac{q(x_k)-z}{q(x_0)-z} < \epsilon$ end.
else
1. Solve (BQP).
2. Let $x_{k+1} = x_k + P(x_k)^{\frac{1}{2}} d'_{x_k}$.
3. Let $k = k + 1$ and return to step 1.

5. **Concluding remarks**. In this paper, we have considered a generalization of a potential-reduction algorithm (which is due to Y.Ye) to a class of nonconvex symmetric programming problems. Many other algorithms based on trust region ideas can be generalized in the similar fashion. This paper is partially based on the work supported by NSF grants DMS0402740 and

DMS0102628.

References

[1] [BS] J.F.Bonnans and A.Shapiro. Perturbation Analysis of Optimization Problems. Springer-Verlag New York, 2000.

[2] [FK] Faraut,J. and A.Koranyi. Analysis on symmetric cones, Oxford University Press, New York, 1994.

[3] [f1] L.Faybusovich. Euclidean Jordan algebras and interior-point algorithms. J. Positivity,331-357,1997.

[4] [f2] L.Faybusovich. A Jordan-algebraic approach to potential-reduction algorithms. Mathematische Zeitchrift, 117-129, 2002.

[5] [f3] L.Faybusovich. Linear systems in Jordan algebras and primal-dual interior-point algorithms, Journal of Comp. and Applied mathematics, vol 86 (1997), pp149-175.

[6] [f4] L.Faybusovich and R. Arana, A long-step Primal-dual algorithm for symmetric programming problem, Syst. Cont.Letter, 43(2001), no1, pp3-7.

[7] [fT] L.Faybusovich and T.Tsuchiya, Primal-dual Algorithms and Infinite-dimensional Jordan algebras of finite rank. math.Progr. 97(2003) no 3 series B, 471-493.

[8] [M] M. Muramatsu, On commutative class of search directions for linear programming over symmetric cone, Journal of optimization Theory and Its Applications, vol 112, no3 (2002), pp 595-625.

[9] [Renegar] Renegar,J. A Mathematical View Of Interior-point Methods in Convex Programming, SIAM Publications, Philadelphia, Pennsylvania, 2001.

[10] [S] S.Smieta and F.Alizadeh, Extensions of primal-dual interior-point algorithms to symmetric cones, Math Prog. 96(2003), no3, series A 409-438.

[11] [T] T.Tsuchiya, A convergence analysis of the scaling-invariant primal-dual path-following algorithms for second-order cone programming, Opt. Methods and Software, 10/11, 141-182, 1999.

[12] [Ye1] Y.Ye. On an affine scaling algorithm for nonconvex quadratic programming. Math.Programming, 52: 285-300,1992.

[13] [Ye2] Y.Ye. Interior Point Algorithms: Theory and Analysis, John Wiley and Sons, New York, 1997.