An interior point algorithm for solving a special QCQP that occurs in bundle methods

Hannes Fendl
Faculty of Mathematics, University of Vienna
Nordbergstraße 15, A-1090 Vienna

February 20, 2012

Abstract. We present an implementation of an interior point algorithm for solving a special QCQP that can be used for determining the search direction in bundle methods for nonlinearly constrained nonsmooth optimization.

Contents

1 Introduction

Let $W, \hat{G} \in \mathbb{R}^{n \times n}_{\text{sym}}$ be positive definite and $|J_k| = m$. We want to derive the interior point framework for the convex QCQP (cf. Fendl [1])

$$\begin{aligned} &\min_{d,v} v + \frac{1}{2} d^T W_p d \\ &\text{s.t. } -\alpha_j + d^T g_j \le v \qquad \text{for } j \in J_k \\ &\quad F - A_j + d^T \hat{g}_j + \tfrac{1}{2} d^T \hat{G} d \le 0 \quad \text{for } j \in J_k \ . \end{aligned} \tag{1}$$

Proposition 1.1 (Optimality conditions of the reduced QCQP)**.** *$(d, v) \in \mathbb{R}^{n+1}$ solves the reduced QCQP (1) if and only if there exists $(\lambda, \mu) \in \mathbb{R}^{2m}_{\ge 0}$ with*

$$\begin{aligned} \Big(W + (\sum_{j \in J_k} \mu_j)\hat{G}\Big)d + \sum_{j \in J_k} \lambda_j g_j + \sum_{j \in J_k} \mu_j \hat{g}_j = o_n& \\ \sum_{j \in J_k} \lambda_j = 1& \\ -\alpha_j + d^T g_j - v \le 0 \text{ for } j \in J_k& \\ F - A_j + d^T \hat{g}_j + \tfrac{1}{2} d^T \hat{G} d \le 0 \text{ for } j \in J_k& \\ \lambda_j(-\alpha_j + d^T g_j - v) = 0 \text{ for } j \in J_k& \\ \mu_j(F - A_j + d^T \hat{g}_j + \tfrac{1}{2} d^T \hat{G} d) = 0 \text{ for } j \in J_k& \ . \end{aligned} \tag{2}$$

Proof. Clear due to the convexity of the reduced QCQP (1). □

2 Interior point method

2.1 Basics

Notation 2.1. We define for $\zeta \in \mathbb{R}$, $z \in \mathbb{R}^m$, and $Z \in \mathbb{R}^{m \times n}$

$$Z + \zeta := \begin{pmatrix} Z_{11} + \zeta & \dots & Z_{1n} + \zeta \\ \vdots & & \vdots \\ Z_{m1} + \zeta & \dots & Z_{mn} + \zeta \end{pmatrix} , \quad Z + z := \left(Z_{:1} + z, \dots, Z_{:n} + z \right) .$$

We write for $\lambda, \mu, s, t \in \mathbb{R}^m$

$$\Lambda := \operatorname{diag}(\lambda) , \quad M := \operatorname{diag}(\mu) , \quad S := \operatorname{diag}(s) , \quad T := \operatorname{diag}(t) . \tag{3}$$

and we define

$$c := \frac{\lambda^T s + \mu^T t}{2m} . \tag{4}$$

2.2 Line search

We define a special neighbourhood for our iterates (cf. Wright [3, p. 166, Equation (8.20)]).

Definition 2.2 (Neighbourhood)**.** Let $\gamma \in (0, 1)$ and $\beta > 0$. We define

$$\begin{aligned} \mathcal{N}_{-\infty}(\gamma, \beta, c) := \Big\{ (d, v, \lambda, \mu, s, t) \in \mathbb{R}^{n+1+4m} : \ \big| \big(W + \operatorname{tr}(M) \hat{G} \big) d + g\lambda + \hat{g}\mu \big|_\infty &\le \beta c , \\ |-1 + \operatorname{tr}(\Lambda)| &\le \beta c , \\ |-\alpha + g^T d - v 1_m + s|_\infty &\le \beta c , \\ |-A + \hat{g}^T d + (F + \tfrac{1}{2} d^T \hat{G} d) 1_m + t|_\infty &\le \beta c , \\ \lambda_j s_j &\ge \gamma c , \\ \mu_j t_j &\ge \gamma c , \\ (\lambda, s) &> 0 , \\ (\mu, t) &> 0 \Big\} . \end{aligned} \tag{5}$$

We want to perform a line search which determines $r \in \{1, \chi, \chi^2, \dots\} \subseteq (0, 1]$, where $\chi \in (0, 1)$ is fixed, such that

$$(d, v, \lambda, \mu, s, t) + r(d^{\text{new}}, v^{\text{new}}, \lambda^{\text{new}}, \mu^{\text{new}}, s^{\text{new}}, t^{\text{new}}) \in \mathcal{N}_{-\infty}(\gamma, \beta, c) \tag{6}$$

$$\frac{(\lambda + r\lambda^{\text{new}})^T (s + r s^{\text{new}}) + (\mu + r\mu^{\text{new}})^T (t + r t^{\text{new}})}{2m} \le (1 - 0.01 r) c . \tag{7}$$

Finally, we make the update

$$(d, v, \lambda, \mu, s, t) := (d, v, \lambda, \mu, s, t) + r(d^{\text{new}}, v^{\text{new}}, \lambda^{\text{new}}, \mu^{\text{new}}, s^{\text{new}}, t^{\text{new}}) ,$$

and iterate this procedure.

Remark 2.3 (Centrality reduction). The left side of the centrality reduction condition (7) is obtained by is obtained by replacing λ, μ, s, and t in the definition of c from (4) with $\lambda + r\lambda^{\text{new}}$, $\mu + r\mu^{\text{new}}$, $s + rs^{\text{new}}$, and $t + rt^{\text{new}}$.

Proposition 2.4 (Line search scaling). *For $r \in \mathbb{R}_{\geq 0}$ the neighbourhood condition (6) is equivalent to*

$$\Big|\Big(\big(W + \mathrm{tr}(M)\hat{G}\big)d + (g\lambda + \hat{g}\mu)\Big) + r\Big(\big(W + \mathrm{tr}(M)\hat{G}\big)d^{\mathrm{new}} + \mathrm{tr}(M^{\mathrm{new}})\hat{G}d + (g\lambda^{\mathrm{new}} + \hat{g}\mu^{\mathrm{new}})\Big) + r^2\mathrm{tr}(M^{\mathrm{new}})\hat{G}d^{\mathrm{new}}\Big|_\infty \leq \beta c \tag{8}$$

$$\big|\big(-1 + \mathrm{tr}(\Lambda)\big) + r \cdot \mathrm{tr}(\Lambda^{\mathrm{new}})\big| \leq \beta c \tag{9}$$

$$|(-\alpha + g^T d - v1_m + s) + r(g^T d^{\mathrm{new}} - v^{\mathrm{new}}1_m + s^{\mathrm{new}})|_\infty \leq \beta c \tag{10}$$

$$\big|\big(-A + \hat{g}^T d + (F + \tfrac{1}{2}d^T\hat{G}d)1_m + t\big) + r\big((\hat{g} + \hat{G}d)^T d^{\mathrm{new}} + t^{\mathrm{new}}\big) + r^2(\tfrac{1}{2}d^{\mathrm{new}^T}\hat{G}d^{\mathrm{new}})1_m\big|_\infty \leq \beta c \tag{11}$$

and

$$(-r)(\lambda_j^{\mathrm{new}}s_j + \lambda_j s_j^{\mathrm{new}}) - r^2\lambda_j^{\mathrm{new}}s_j^{\mathrm{new}} \leq \lambda_j s_j - \gamma c \tag{12}$$
$$(-r)(\mu_j^{\mathrm{new}}t_j + \mu_j t_j^{\mathrm{new}}) - r^2\mu_j^{\mathrm{new}}t_j^{\mathrm{new}} \leq \mu_j t_j - \gamma c \tag{13}$$

and

$$(-r)(\lambda^{\mathrm{new}}, s^{\mathrm{new}}) < (\lambda, s) \tag{14}$$
$$(-r)(\mu^{\mathrm{new}}, t^{\mathrm{new}}) < (\mu, t)\ . \tag{15}$$

For $r \in \mathbb{R}_{>0}$ the centrality reduction condition (7) is equivalent to

$$0 \geq \Big(0.01c + \frac{\lambda^{\mathrm{new}^T}s + \lambda^T s^{\mathrm{new}} + \mu^{\mathrm{new}^T}t + \mu^T t^{\mathrm{new}}}{2m}\Big) + r\frac{\lambda^{\mathrm{new}^T}s^{\mathrm{new}} + \mu^{\mathrm{new}^T}t^{\mathrm{new}}}{2m}\ . \tag{16}$$

Proof. Inserting $d + rd^{\mathrm{new}}$, $\lambda + r\lambda^{\mathrm{new}}$, and $\mu + r\mu^{\mathrm{new}}$ into the error estimation of the gradient of the Lagrangian from (5) yields

$$\begin{aligned}
\beta c &\geq \big|\big(W + \mathrm{tr}(M + rM^{\mathrm{new}})\hat{G}\big)(d + rd^{\mathrm{new}}) + g(\lambda + r\lambda^{\mathrm{new}}) + \hat{g}(\mu + r\mu^{\mathrm{new}})\big|_\infty \\
&= \Big|\Big(\big(W + \mathrm{tr}(M)\hat{G}\big) + r \cdot \mathrm{tr}(M^{\mathrm{new}})\hat{G}\Big)(d + rd^{\mathrm{new}}) + (g\lambda + \hat{g}\mu) + r(g\lambda^{\mathrm{new}} + \hat{g}\mu^{\mathrm{new}})\Big|_\infty \\
&= \big|\big(W + \mathrm{tr}(M)\hat{G}\big)(d + rd^{\mathrm{new}}) + r \cdot \mathrm{tr}(M^{\mathrm{new}})\hat{G}(d + rd^{\mathrm{new}}) + (g\lambda + \hat{g}\mu) + r(g\lambda^{\mathrm{new}} + \hat{g}\mu^{\mathrm{new}})\big|_\infty \\
&= \big|\big(W + \mathrm{tr}(M)\hat{G}\big)d + r\big(W + \mathrm{tr}(M)\hat{G}\big)d^{\mathrm{new}} + r \cdot \mathrm{tr}(M^{\mathrm{new}})\hat{G}d + r^2\mathrm{tr}(M^{\mathrm{new}})\hat{G}d^{\mathrm{new}} \\
&\quad + (g\lambda + \hat{g}\mu) + r(g\lambda^{\mathrm{new}} + \hat{g}\mu^{\mathrm{new}})\big|_\infty \\
&= \Big|\Big(\big(W + \mathrm{tr}(M)\hat{G}\big)d + (g\lambda + \hat{g}\mu)\Big) + r\Big(\big(W + \mathrm{tr}(M)\hat{G}\big)d^{\mathrm{new}} + \mathrm{tr}(M^{\mathrm{new}})\hat{G}d + (g\lambda^{\mathrm{new}} + \hat{g}\mu^{\mathrm{new}})\Big) \\
&\quad + r^2\mathrm{tr}(M^{\mathrm{new}})\hat{G}d^{\mathrm{new}}\Big|_\infty\ .
\end{aligned}$$

Inserting $\lambda + r\lambda^{\mathrm{new}}$ into the feasibility estimation of the convex combination constraint from (5) yields

$$\beta c \geq |-1 + \mathrm{tr}(\Lambda + r\Lambda^{\mathrm{new}})| = \big|\big(-1 + \mathrm{tr}(\Lambda)\big) + r \cdot \mathrm{tr}(\Lambda^{\mathrm{new}})\big|\ .$$

Inserting $d + rd^{\mathrm{new}}$, $v + rv^{\mathrm{new}}$, and $s + rs^{\mathrm{new}}$ into the feasibility estimations of the linear inequality constraints from (5) yields

$$\begin{aligned}
\beta c &\geq |-\alpha + g^T(d + rd^{\mathrm{new}}) - (v + rv^{\mathrm{new}})1_m + (s + rs^{\mathrm{new}})|_\infty \\
&= |(-\alpha + g^T d - v1_m + s) + r(g^T d^{\mathrm{new}} - v^{\mathrm{new}}1_m + s^{\mathrm{new}})|_\infty\ .
\end{aligned}$$

Since we have

$$
\begin{aligned}
\tfrac{1}{2}(d+rd^{\text{new}})^T\hat{G}(d+rd^{\text{new}}) &= \tfrac{1}{2}(d^T\hat{G}d + rd^{\text{new}^T}\hat{G}d + rd^T\hat{G}d^{\text{new}} + r^2 d^{\text{new}^T}\hat{G}d^{\text{new}}) \\
&= \tfrac{1}{2}\big(d^T\hat{G}d + r(d^{\text{new}^T}\hat{G}d)^T + rd^T\hat{G}d^{\text{new}} + r^2 d^{\text{new}^T}\hat{G}d^{\text{new}}\big) \\
&= \tfrac{1}{2}\big(d^T\hat{G}d + rd^T\hat{G}^T d^{\text{new}} + rd^T\hat{G}d^{\text{new}} + r^2 d^{\text{new}^T}\hat{G}d^{\text{new}}\big) \\
&\overset{\substack{\hat{G}\in\mathbb{R}^{n\times n}_{\text{sym}} \\ \downarrow}}{=} \tfrac{1}{2}(d^T\hat{G}d + rd^T\hat{G}d^{\text{new}} + rd^T\hat{G}d^{\text{new}} + r^2 d^{\text{new}^T}\hat{G}d^{\text{new}}) \\
&= \tfrac{1}{2}(d^T\hat{G}d + 2rd^T\hat{G}d^{\text{new}} + r^2 d^{\text{new}^T}\hat{G}d^{\text{new}}) \\
&= \tfrac{1}{2}d^T\hat{G}d + r(d^T\hat{G})d^{\text{new}} + \tfrac{1}{2}r^2 d^{\text{new}^T}\hat{G}d^{\text{new}} \ , \qquad (17)
\end{aligned}
$$

inserting $d+rd^{\text{new}}$ and $t+rt^{\text{new}}$ into the feasibility estimations of the quadratic inequality constraints from (5) yields

$$
\begin{aligned}
\beta c &\geq |-A + \hat{g}^T(d+rd^{\text{new}}) + \big(F + \tfrac{1}{2}(d+rd^{\text{new}})^T\hat{G}(d+rd^{\text{new}})\big)1_m + (t+rt^{\text{new}})|_\infty \\
&= |(-A+\hat{g}^Td+t) + r(\hat{g}^Td^{\text{new}}+t^{\text{new}}) + \big(F + \tfrac{1}{2}(d+rd^{\text{new}})^T\hat{G}(d+rd^{\text{new}})\big)1_m|_\infty \\
&\overset{(17)}{=} \big|(-A+\hat{g}^Td+t) + r(\hat{g}^Td^{\text{new}}+t^{\text{new}}) + \big(F + \tfrac{1}{2}d^T\hat{G}d + r(d^T\hat{G})d^{\text{new}} + \tfrac{1}{2}r^2 d^{\text{new}^T}\hat{G}d^{\text{new}}\big)1_m\big|_\infty \\
&= \big|\big(-A+\hat{g}^Td + (F+\tfrac{1}{2}d^T\hat{G}d)1_m + t\big) + r\big(\hat{g}^Td^{\text{new}} + (d^T\hat{G})d^{\text{new}}1_m + t^{\text{new}}\big) + r^2(\tfrac{1}{2}d^{\text{new}^T}\hat{G}d^{\text{new}})1_m\big|_\infty \\
&= \big|\big(-A+\hat{g}^Td + (F+\tfrac{1}{2}d^T\hat{G}d)1_m + t\big) + r\big((\hat{g}^T + d^T\hat{G})d^{\text{new}} + t^{\text{new}}\big) + r^2(\tfrac{1}{2}d^{\text{new}^T}\hat{G}d^{\text{new}})1_m\big|_\infty \\
&\overset{\substack{\hat{G}\in\mathbb{R}^{n\times n}_{\text{sym}} \\ \downarrow}}{=} \big|\big(-A+\hat{g}^Td + (F+\tfrac{1}{2}d^T\hat{G}d)1_m + t\big) + r\big((\hat{g}^T + d^T\hat{G}^T)d^{\text{new}} + t^{\text{new}}\big) + r^2(\tfrac{1}{2}d^{\text{new}^T}\hat{G}d^{\text{new}})1_m\big|_\infty \\
&= \big|\big(-A+\hat{g}^Td + (F+\tfrac{1}{2}d^T\hat{G}d)1_m + t\big) + r\big((\hat{g} + \hat{G}d)^Td^{\text{new}} + t^{\text{new}}\big) + r^2(\tfrac{1}{2}d^{\text{new}^T}\hat{G}d^{\text{new}})1_m\big|_\infty \ .
\end{aligned}
$$

Inserting $\lambda + r\lambda^{\text{new}}$ and $s + rs^{\text{new}}$ into the estimations for the complementarity conditions from (5) yields

$$
\begin{aligned}
\gamma c &\leq (\lambda_j + r\lambda_j^{\text{new}})(s_j + rs_j^{\text{new}}) \\
&= \lambda_j s_j + r\lambda_j^{\text{new}}s_j + r\lambda_j s_j^{\text{new}} + r^2\lambda_j^{\text{new}}s_j^{\text{new}} \\
&= \lambda_j s_j + r(\lambda_j^{\text{new}}s_j + \lambda_j s_j^{\text{new}}) + r^2\lambda_j^{\text{new}}s_j^{\text{new}} \\
\iff (-r)(\lambda_j^{\text{new}}s_j + \lambda_j s_j^{\text{new}}) - r^2\lambda_j^{\text{new}}s_j^{\text{new}} &\leq \lambda_j s_j - \gamma c
\end{aligned}
$$

and

$$
\begin{aligned}
\gamma c &\leq (\mu_j + r\mu_j^{\text{new}})(t_j + rt_j^{\text{new}}) \\
&= \mu_j t_j + r\mu_j^{\text{new}}t_j + r\mu_j t_j^{\text{new}} + r^2\mu_j^{\text{new}}t_j^{\text{new}} \\
&= \mu_j t_j + r(\mu_j^{\text{new}}t_j + \mu_j t_j^{\text{new}}) + r^2\mu_j^{\text{new}}t_j^{\text{new}} \\
\iff (-r)(\mu_j^{\text{new}}t_j + \mu_j t_j^{\text{new}}) - r^2\mu_j^{\text{new}}t_j^{\text{new}} &\leq \mu_j t_j - \gamma c \ .
\end{aligned}
$$

Inserting $\lambda + r\lambda^{\text{new}}$ and $s + rs^{\text{new}}$ into the positivity conditions from (5) yields

$$0 < (\lambda + r\lambda^{\text{new}}, s + rs^{\text{new}}) = (\lambda, s) + r(\lambda^{\text{new}}, s^{\text{new}}) \iff (-r)(\lambda^{\text{new}}, s^{\text{new}}) < (\lambda, s)$$

and

$$0 < (\mu + r\mu^{\text{new}}, t + rt^{\text{new}}) = (\mu, t) + r(\mu^{\text{new}}, t^{\text{new}}) \iff (-r)(\mu^{\text{new}}, t^{\text{new}}) < (\mu, t) \ .$$

Finally, we obtain for the centrality reduction condition (7)

$$
\begin{aligned}
c - 0.01rc &= (1-0.01r)c \\
&\overset{(7)}{\geq} \frac{(\lambda + r\lambda^{\text{new}})^T(s + rs^{\text{new}}) + (\mu + r\mu^{\text{new}})^T(t + rt^{\text{new}})}{2m} \\
&= \frac{(\lambda^T s + r{\lambda^{\text{new}}}^T s + r\lambda^T s^{\text{new}} + r^2 {\lambda^{\text{new}}}^T s^{\text{new}}) + (\mu^T s + r{\mu^{\text{new}}}^T t + r\mu^T t^{\text{new}} + r^2 {\mu^{\text{new}}}^T t^{\text{new}})}{2m} \\
&= \frac{(\lambda^T s + \mu^T s) + r({\lambda^{\text{new}}}^T s + \lambda^T s^{\text{new}} + {\mu^{\text{new}}}^T t + \mu^T t^{\text{new}}) + r^2({\lambda^{\text{new}}}^T s^{\text{new}} + {\mu^{\text{new}}}^T t^{\text{new}})}{2m} \\
&\overset{(4)}{=} c + \frac{r({\lambda^{\text{new}}}^T s + \lambda^T s^{\text{new}} + {\mu^{\text{new}}}^T t + \mu^T t^{\text{new}}) + r^2({\lambda^{\text{new}}}^T s^{\text{new}} + {\mu^{\text{new}}}^T t^{\text{new}})}{2m} \ ,
\end{aligned}
$$

which is—after substracting c on both sides—equivalent to

$$
-0.01rc \geq \frac{r({\lambda^{\text{new}}}^T s + \lambda^T s^{\text{new}} + {\mu^{\text{new}}}^T t + \mu^T t^{\text{new}}) + r^2({\lambda^{\text{new}}}^T s^{\text{new}} + {\mu^{\text{new}}}^T t^{\text{new}})}{2m}
$$

and therefore dividing by $r \in (0,1]$ yields

$$
\begin{aligned}
-0.01c &\geq \frac{({\lambda^{\text{new}}}^T s + \lambda^T s^{\text{new}} + {\mu^{\text{new}}}^T t + \mu^T t^{\text{new}}) + r({\lambda^{\text{new}}}^T s^{\text{new}} + {\mu^{\text{new}}}^T t^{\text{new}})}{2m} \\
&= \frac{{\lambda^{\text{new}}}^T s + \lambda^T s^{\text{new}} + {\mu^{\text{new}}}^T t + \mu^T t^{\text{new}}}{2m} + r\frac{{\lambda^{\text{new}}}^T s^{\text{new}} + {\mu^{\text{new}}}^T t^{\text{new}}}{2m}
\end{aligned}
$$

which is equivalent to

$$
0 \geq \left(0.01c + \frac{{\lambda^{\text{new}}}^T s + \lambda^T s^{\text{new}} + {\mu^{\text{new}}}^T t + \mu^T t^{\text{new}}}{2m}\right) + r\frac{{\lambda^{\text{new}}}^T s^{\text{new}} + {\mu^{\text{new}}}^T t^{\text{new}}}{2m} \ . \qquad \square
$$

Proposition 2.5 (Starting point)**.** *If we set*

$$
d := o_n \ , \quad v := 0 \ , \quad \lambda := 1_m \ , \quad \mu := 1_m \ , \quad s := 1_m \ , \quad t := 1_m \tag{18}
$$

and if we choose $\gamma \in (0,1)$ *and*

$$
\beta \geq \max\left(|(g+\hat{g})1_m|_\infty, |-1+m|, |-\alpha + 1_m|, |-A + (1+F)1_m|\right) \tag{19}
$$

with $\beta > 0$*, then*

$$
(d, v, \lambda, \mu, s, t) \in \mathcal{N}_{-\infty}(\gamma, \beta, c) \ .
$$

Proof. We calculate

$$
c \overset{(4)}{=} \frac{\lambda^T s + \mu^T t}{2m} \overset{(18)}{=} \frac{1_m^T 1_m + 1_m^T 1_m}{2m} = \frac{m+m}{2m} = 1 \ . \tag{20}
$$

Now we verify the conditions in the definition of $\mathcal{N}_{-\infty}(\gamma, \beta, c)$ from (5): We obtain

$$
\left|\left(W + \operatorname{tr}(M)\hat{G}\right)d + g\lambda + \hat{g}\mu\right|_\infty \overset{(18)}{=} |g1_m + \hat{g}1_m|_\infty \overset{(19)}{\leq} \beta \overset{(20)}{=} \beta c
$$

and

$$
|-1 + \operatorname{tr}(\Lambda)| \overset{(18)}{=} |-1+m| \overset{(19)}{\leq} \beta \overset{(20)}{=} \beta c
$$

and

$$
|-\alpha + g^T d - v1_m + s|_\infty \overset{(18)}{=} |-\alpha + 1_m|_\infty \overset{(19)}{\leq} \beta \overset{(20)}{=} \beta c
$$

and

$$|-A + \hat{g}^T d + (F + \tfrac{1}{2} d^T \hat{G} d) 1_m + t|_\infty \overset{(18)}{=} |-A + F 1_m + 1_m|_\infty \overset{(19)}{\leq} \beta \overset{(20)}{=} \beta c \;.$$

Furthermore, we have

$$\lambda_j s_j \overset{(18)}{=} 1 \overset{\substack{\gamma \in (0,1) \\ \downarrow}}{>} \gamma \overset{(20)}{=} \gamma c \;, \quad \mu_j t_j \overset{(18)}{=} 1 \overset{\substack{\gamma \in (0,1) \\ \downarrow}}{>} \gamma \overset{(20)}{=} \gamma c \;.$$

Finally, we show the positivity condition

$$(\lambda, s) \overset{(18)}{=} (1_m, 1_m) > 0 \;, \quad (\mu, t) \overset{(18)}{=} (1_m, 1_m) > 0 \;. \qquad \square$$

2.3 Search direction

Proposition 2.6 (Variant of the optimality conditions). *$(d, v) \in \mathbb{R}^{n+1}$ solves the reduced QCQP (1) if and only if there exists $(\lambda, \mu, s, t) \in \mathbb{R}^{4m}$ such that*

$$\begin{aligned}
\Big(W + (\sum_{j \in J_k} \mu_j) \hat{G}\Big) d + \sum_{j \in J_k} \lambda_j g_j + \mu_j \hat{g}_j &= o_n \\
\sum_{j \in J_k} \lambda_j &= 1 \\
-\alpha_j + d^T g_j - v + s_j &= 0 \text{ for } j \in J_k \\
F - A_j + d^T \hat{g}_j + \tfrac{1}{2} d^T \hat{G} d + t_j &= 0 \text{ for } j \in J_k \\
\lambda_j s_j &= 0 \text{ for } j \in J_k \\
\mu_j t_j &= 0 \text{ for } j \in J_k \\
(\lambda, s) &\geq o_{2m} \\
(\mu, t) &\geq o_{2m} \;.
\end{aligned} \tag{21}$$

Proof. Since $(d, v) \in \mathbb{R}^{n+1}$ solves the reduced QCQP (1) if and only if (2) holds, we obtain

$$\begin{aligned}
-\alpha_j + d^T g_j - v \leq 0 &\iff -\alpha_j + d^T g_j - v + s_j = 0 \quad \wedge \quad s_j \geq 0 \\
F - A_j + d^T \hat{g}_j + \tfrac{1}{2} d^T \hat{G} d \leq 0 &\iff F - A_j + d^T \hat{g}_j + \tfrac{1}{2} d^T \hat{G} d + t_j = 0 \quad \wedge \quad t_j \geq 0 \;,
\end{aligned}$$

adding $\lambda_j s_j$ resp. $\mu_j t_j$ on both sides of the complementarity conditions makes them equivalent to

$$\begin{aligned}
\lambda_j(-\alpha_j + d^T g_j - v) + \lambda_j s_j = \lambda_j s_j &\iff \lambda_j s_j = 0 \\
\mu_j(F - A_j + d^T \hat{g}_j + \tfrac{1}{2} d^T \hat{G} d) + \mu_j t_j = \mu_j t_j &\iff \mu_j t_j = 0 \;,
\end{aligned}$$

which yields the desired result. $\square$

Proposition 2.7 (Trace). *Let $D \in \mathbb{R}^{m \times m}_{\text{diag}}$. Then*

$$1_m^T D 1_m = \operatorname{tr}(D) \;. \tag{22}$$

Proof. We calculate

$$1_m^T D 1_m = 1_m^T \begin{pmatrix} D_{11} \\ \vdots \\ D_{mm} \end{pmatrix} = \sum_{j=1}^m D_{jj} = \operatorname{tr}(D) \;. \qquad \square$$

An interior point method applies Newton's method to an appropriately relaxed version of (21) to obtain the linear system, whose solution is used as the search direction (cf. Wright [3, p. 165, Equation (8.18) & (8.19), and p. 210, Three Forms of the Step Equation]).

Proposition 2.8 (Search direction)**.** *If we apply Newton's method to the following relaxed version of (21)*

$$\begin{aligned}
\Big(W + (\sum_{j\in J_k} \mu_j)\hat{G}\Big)d + \sum_{j\in J_k} \lambda_j g_j + \mu_j \hat{g}_j &= o_n \\
\sum_{j\in J_k} \lambda_j &= 1 \\
-\alpha_j + d^T g_j - v + s_j &= 0 \;\text{ for } j \in J_k \\
F - A_j + d^T \hat{g}_j + \tfrac{1}{2} d^T \hat{G} d + t_j &= 0 \;\text{ for } j \in J_k \\
\lambda_j s_j &= \sigma c \text{ for } j \in J_k \\
\mu_j t_j &= \sigma c \text{ for } j \in J_k \ ,
\end{aligned} \tag{23}$$

then we obtain a linear system that can be expressed in one of the following six forms, where $\sigma \in [\sigma_{\min}, \sigma_{\max}]$ *with* $0 < \sigma_{\min} < \sigma_{\max} < 1$ *and we use the abbreviation*

$$x := (d^{\text{new}}, v^{\text{new}}, \lambda^{\text{new}}, \mu^{\text{new}}, s^{\text{new}}, t^{\text{new}}) \in \mathbb{R}^{n+1+4m} \ . \tag{24}$$

1. *The general form of the linear system reads*

$$\bar{B}x = \bar{b} \ ,$$

where

$$\bar{B} := \begin{pmatrix} W + \mathrm{tr}(M)\hat{G} & o_n & g & \hat{g} + \hat{G}d & O_{n\times m} & O_{n\times m} \\ o_n^T & 0 & -1_m^T & o_m^T & o_m^T & o_m^T \\ g^T & -1_m & O_{m\times m} & O_{m\times m} & I_m & O_{m\times m} \\ (\hat{g} + \hat{G}d)^T & o_m & O_{m\times m} & O_{m\times m} & O_{m\times m} & I_m \\ O_{m\times n} & o_m & S & O_{m\times m} & \Lambda & O_{m\times m} \\ O_{m\times n} & o_m & O_{m\times m} & T & O_{m\times m} & M \end{pmatrix} \in \mathbb{R}^{(n+1+4m)\times(n+1+4m)} \ ,$$

$$\bar{b} := -\begin{pmatrix} \big(W + \mathrm{tr}(M)\hat{G}\big)d + g\lambda + \hat{g}\mu \\ 1 - \mathrm{tr}(\Lambda) \\ -\alpha + g^T d - v 1_m + s \\ -A + \hat{g}^T d + (F + \frac{1}{2} d^T \hat{G} d)1_m + t \\ (\Lambda S - \sigma c I_m)1_m \\ (MT - \sigma c I_m)1_m \end{pmatrix} \in \mathbb{R}^{n+1+4m} \ .$$

2. *The symmetric form of the linear system reads*

$$\tilde{B}x = \tilde{b} \ ,$$

where

$$\tilde{B} := \begin{pmatrix} W + \mathrm{tr}(M)\hat{G} & o_n & g & \hat{g} + \hat{G}d & O_{n\times m} & O_{n\times m} \\ o_n^T & 0 & -1_m^T & o_m^T & o_m^T & o_m^T \\ g^T & -1_m & O_{m\times m} & O_{m\times m} & I_m & O_{m\times m} \\ (\hat{g} + \hat{G}d)^T & o_m & O_{m\times m} & O_{m\times m} & O_{m\times m} & I_m \\ O_{m\times n} & o_m & I_m & O_{m\times m} & S^{-1}\Lambda & O_{m\times m} \\ O_{m\times n} & o_m & O_{m\times m} & I_m & O_{m\times m} & T^{-1}M \end{pmatrix} \in \mathbb{R}^{(n+1+4m)\times(n+1+4m)}_{\text{sym}} \ ,$$

$$\tilde{b} := -\begin{pmatrix} \big(W + \mathrm{tr}(M)\hat{G}\big)d + g\lambda + \hat{g}\mu \\ 1 - \mathrm{tr}(\Lambda) \\ -\alpha + g^T d - v 1_m + s \\ -A + \hat{g}^T d + (F + \frac{1}{2} d^T \hat{G} d)1_m + t \\ (\Lambda - \sigma c S^{-1})1_m \\ (M - \sigma c T^{-1})1_m \end{pmatrix} \in \mathbb{R}^{n+1+4m} \ .$$

3. *The compact form of the linear system reads*

$$\hat{B}x_{1:(n+1+2m)} = \hat{b} \ ,$$

where

$$\hat{B} := \begin{pmatrix} W + \mathrm{tr}(M)\hat{G} & o_n & g & \hat{g} + \hat{G}d \\ o_n^T & 0 & -1_m^T & o_m^T \\ g^T & -1_m & -\Lambda^{-1}S & O_{m\times m} \\ (\hat{g} + \hat{G}d)^T & o_m & O_{m\times m} & -M^{-1}T \end{pmatrix} \in \mathbb{R}^{(n+1+2m)\times(n+1+2m)}_{\mathrm{sym}} \ ,$$

$$\hat{b} := -\begin{pmatrix} (W + \mathrm{tr}(M)\hat{G})d + g\lambda + \hat{g}\mu \\ 1 - \mathrm{tr}(\Lambda) \\ -\alpha + g^T d - v1_m + \sigma c\Lambda^{-1}1_m \\ -A + \hat{g}^T d + (F + \frac{1}{2}d^T\hat{G}d)1_m + \sigma cM^{-1}1_m \end{pmatrix} \in \mathbb{R}^{n+1+2m} \ ,$$

and s^{new} and t^{new} can be computed by

$$\begin{aligned} s^{\mathrm{new}} &= -\Lambda^{-1}S\lambda^{\mathrm{new}} - (S - \sigma c\Lambda^{-1})1_m \\ t^{\mathrm{new}} &= -M^{-1}T\mu^{\mathrm{new}} - (T - \sigma cM^{-1})1_m \ . \end{aligned} \tag{25}$$

4. *The practical form of the linear system reads*

$$\check{B}x_{1:(n+1)} = \check{b} \ ,$$

where

$$\check{B} := \begin{pmatrix} (W + \mathrm{tr}(M)\hat{G}) + g\Lambda S^{-1}g^T + (\hat{g} + \hat{G}d)MT^{-1}(\hat{g} + \hat{G}d)^T & -g\Lambda S^{-1}1_m \\ -(g\Lambda S^{-1}1_m)^T & \mathrm{tr}(\Lambda S^{-1}) \end{pmatrix} \in \mathbb{R}^{(n+1)\times(n+1)}_{\mathrm{sym}}$$

$$\begin{aligned} \check{b}_{1:n} := &-\Big(\big(W + \mathrm{tr}(M)\hat{G}\big)d + g\lambda + \hat{g}\mu\Big) \\ &- \big(g\Lambda S^{-1}(-\alpha + g^T d - v1_m) + \sigma cgS^{-1}1_m\big) \\ &- \Big((\hat{g} + \hat{G}d)MT^{-1}\big(-A + \hat{g}^T d + (F + \tfrac{1}{2}d^T\hat{G}d)1_m\big) + \sigma c(\hat{g} + \hat{G}d)T^{-1}1_m\Big) \in \mathbb{R}^n \end{aligned}$$

$$\check{b}_{n+1} = -\big(1 - \mathrm{tr}(\Lambda)\big) + 1_m^T\Lambda S^{-1}(-\alpha + g^T d) - v\cdot\mathrm{tr}(\Lambda S^{-1}) + \sigma c\cdot\mathrm{tr}(S^{-1}) \in \mathbb{R} \ ,$$

and λ^{new} and μ^{new} can be computed by

$$\begin{aligned} \lambda^{\mathrm{new}} &= \Lambda S^{-1}(g^T, -1_m)\binom{d^{\mathrm{new}}}{v^{\mathrm{new}}} + \Lambda S^{-1}(-\alpha + g^T d - v1_m) + \sigma cS^{-1}1_m \\ \mu^{\mathrm{new}} &= MT^{-1}(\hat{g} + \hat{G}d)^T d^{\mathrm{new}} + MT^{-1}\big(-A + \hat{g}^T d + (F + \tfrac{1}{2}d^T\hat{G}d)1_m\big) + \sigma cT^{-1}1_m \ , \end{aligned} \tag{26}$$

and s^{new} and t^{new} can be computed according to (25).

5. *The best form of the linear system reads*

$$\vec{B}x_{1:n} = \vec{b} \ ,$$

where

$$\vec{B} := W + \mathrm{tr}(M)\hat{G} + g\Lambda S^{-1}g^T + (\hat{g} + \hat{G}d)MT^{-1}(\hat{g} + \hat{G}d)^T - \frac{(g\Lambda S^{-1}1_m)(g\Lambda S^{-1}1_m)^T}{\mathrm{tr}(\Lambda S^{-1})} \in \mathbb{R}^{n\times n}_{\mathrm{sym}}$$

$$\begin{aligned} \vec{b} := &-\Big(\big(W + \mathrm{tr}(M)\hat{G}\big)d + g\lambda + \hat{g}\mu\Big) \\ &- \big(g\Lambda S^{-1}(-\alpha + g^T d - v1_m) + \sigma cgS^{-1}1_m\big) \\ &- \Big((\hat{g} + \hat{G}d)MT^{-1}\big(-A + \hat{g}^T d + (F + \tfrac{1}{2}d^T\hat{G}d)1_m\big) + \sigma c(\hat{g} + \hat{G}d)T^{-1}1_m\Big) \\ &+ g\Lambda S^{-1}1_m\Big(\tfrac{-(1-\mathrm{tr}(\Lambda)) + 1_m^T\Lambda S^{-1}(-\alpha + g^T d) - v\cdot\mathrm{tr}(\Lambda S^{-1}) + \sigma c\cdot\mathrm{tr}(S^{-1})}{\mathrm{tr}(\Lambda S^{-1})}\Big) \in \mathbb{R}^n \ , \end{aligned}$$

and v^{new} can be computed by

$$v^{\mathrm{new}} = \frac{(g\Lambda S^{-1}1_m)^T d^{\mathrm{new}}}{\mathrm{tr}(\Lambda S^{-1})} + \frac{-(1-\mathrm{tr}(\Lambda)) + 1_m^T\Lambda S^{-1}(-\alpha + g^T d) - v\cdot\mathrm{tr}(\Lambda S^{-1}) + \sigma c\cdot\mathrm{tr}(S^{-1})}{\mathrm{tr}(\Lambda S^{-1})} \ ,$$

and λ^{new} and μ^{new} can be computed according to (26), and s^{new} and t^{new} can be computed according to (25).

6. *The best positive definite form of the linear system reads: Set*

$$z_1 := -\big(1-\operatorname{tr}(\Lambda)\big) + 1_m^T \Lambda S^{-1}(-\alpha + g^T d) - v \cdot \operatorname{tr}(\Lambda S^{-1}) + \sigma c \cdot \operatorname{tr}(S^{-1}) \in \mathbb{R} \tag{27}$$

and

$$\begin{aligned} z_2 := &-(g\Lambda S^{-1} 1_m)^T \big(W + \operatorname{tr}(M)\hat{G} + g\Lambda S^{-1} g^T + (\hat{g}+\hat{G}d)MT^{-1}(\hat{g}+\hat{G}d)^T\big)^{-1} \cdot \\ &\cdot \Bigg(-\Big(\big(W + \operatorname{tr}(M)\hat{G}\big)d + g\lambda + \hat{g}\mu\Big) \\ &\quad - \big(g\Lambda S^{-1}(-\alpha + g^T d - v 1_m) + \sigma c g S^{-1} 1_m\big) \\ &\quad - \Big((\hat{g}+\hat{G}d)MT^{-1}\big(-A + \hat{g}^T d + (F + \tfrac{1}{2} d^T \hat{G} d)1_m\big) + \sigma c(\hat{g}+\hat{G}d)T^{-1}1_m\Big)\Bigg) \in \mathbb{R} \end{aligned} \tag{28}$$

and

$$\begin{aligned} z_3 := &\\ &\operatorname{tr}(\Lambda S^{-1}) - (g\Lambda S^{-1}1_m)^T\big(W + \operatorname{tr}(M)\hat{G} + g\Lambda S^{-1}g^T + (\hat{g}+\hat{G}d)MT^{-1}(\hat{g}+\hat{G}d)^T\big)^{-1}(g\Lambda S^{-1}1_m) \in \mathbb{R} \end{aligned} \tag{29}$$

and compute

$$v^{\text{new}} = \tfrac{z_1 - z_2}{z_3} \in \mathbb{R} \ . \tag{30}$$

Solve the symmetric positive definite linear system

$$B x_{1:n} = b \ , \tag{31}$$

where

$$B := W + \operatorname{tr}(M)\hat{G} + g\Lambda S^{-1}g^T + (\hat{g}+\hat{G}d)MT^{-1}(\hat{g}+\hat{G}d)^T \in \mathbb{R}^{n\times n}_{\text{sym}} \text{ positive definite} \tag{32}$$

$$\begin{aligned} b := &-\Big(\big(W + \operatorname{tr}(M)\hat{G}\big)d + g\lambda + \hat{g}\mu\Big) \\ &- \big(g\Lambda S^{-1}(-\alpha + g^T d - v1_m) + \sigma c g S^{-1}1_m\big) \\ &- \Big((\hat{g}+\hat{G}d)MT^{-1}\big(-A + \hat{g}^T d + (F + \tfrac{1}{2}d^T\hat{G}d)1_m\big) + \sigma c(\hat{g}+\hat{G}d)T^{-1}1_m\Big) \\ &+ g\Lambda S^{-1}1_m v^{\text{new}} \in \mathbb{R}^n \ , \end{aligned} \tag{33}$$

and compute λ^{new} *and* μ^{new} *according to (26), and compute* s^{new} *and* t^{new} *according to (25).*

Proof. We start by calculating the necessary linearizations: Linearizing the first equation of (23) yields

$$\begin{aligned} o_n &= \Big(W + (\sum_{j\in J_k}\mu_j)\hat{G}\Big)d + \sum_{j\in J_k}\lambda_j g_j + \mu_j\hat{g}_j + \nabla_{(d,\lambda,\mu)}\bigg(\Big(W + (\sum_{j\in J_k}\mu_j)\hat{G}\Big)d + \sum_{j\in J_k}\lambda_j g_j + \mu_j \hat{g}_j\bigg)\begin{pmatrix} d^{\text{new}} \\ \lambda^{\text{new}} \\ \mu^{\text{new}}\end{pmatrix} \\ &= \Big(W + (\sum_{j\in J_k}\mu_j)\hat{G}\Big)d + \sum_{j\in J_k}\lambda_j g_j + \mu_j\hat{g}_j + \Big(W + (\sum_{j\in J_k}\mu_j)\hat{G}, g, \hat{g}+\hat{G}d\Big)\begin{pmatrix} d^{\text{new}} \\ \lambda^{\text{new}} \\ \mu^{\text{new}}\end{pmatrix} \\ &\overset{(22),(3)}{=} \big(W + \operatorname{tr}(M)\hat{G}\big)d + g\lambda + \hat{g}\mu + \big(W + \operatorname{tr}(M)\hat{G}, g, \hat{g}+\hat{G}d\big)\begin{pmatrix} d^{\text{new}} \\ \lambda^{\text{new}} \\ \mu^{\text{new}}\end{pmatrix} \ , \end{aligned}$$

which is equivalent to

$$-\Big(\big(W+\mathrm{tr}(M)\hat{G}\big)d+g\lambda+\hat{g}\mu\Big)=\big(W+\mathrm{tr}(M)\hat{G},g,\hat{g}+\hat{G}d\big)\begin{pmatrix}d^{\mathrm{new}}\\ \lambda^{\mathrm{new}}\\ \mu^{\mathrm{new}}\end{pmatrix}\ . \tag{34}$$

Linearizing the second equation of (23) yields

$$\begin{aligned}1&=\sum_{j\in J_k}\lambda_j+\nabla_\lambda\Big(\sum_{j\in J_k}\lambda_j\Big)\lambda^{\mathrm{new}}\\ &=\sum_{j\in J_k}\lambda_j+1_m^T\lambda^{\mathrm{new}}\\ &\overset{(22),(3)}{\downarrow \atop =}\ \mathrm{tr}(\Lambda)+1_m^T\lambda^{\mathrm{new}}\\ \iff\ (-1_m)^T\lambda^{\mathrm{new}}&=-1+\mathrm{tr}(\Lambda)\\ &=-\big(1-\mathrm{tr}(\Lambda)\big)\ . \end{aligned}\tag{35}$$

Linearizing the third equation of (23) yields

$$\begin{aligned}0&=-\alpha_j+d^Tg_j-v+s_j+\nabla_{(d,v,s_j)}(-\alpha_j+d^Tg_j-v+s_j)\begin{pmatrix}d^{\mathrm{new}}\\ v^{\mathrm{new}}\\ s_j^{\mathrm{new}}\end{pmatrix}\\ &=-\alpha_j+d^Tg_j-v+s_j+(g_j^T,-1,1)\begin{pmatrix}d^{\mathrm{new}}\\ v^{\mathrm{new}}\\ s_j^{\mathrm{new}}\end{pmatrix}\quad\text{for all } j\in J_k\\ \iff\ o_m&=-\alpha+g^Td-v1_m+s+(g^T,-1_m,I_m)\begin{pmatrix}d^{\mathrm{new}}\\ v^{\mathrm{new}}\\ s^{\mathrm{new}}\end{pmatrix}\ ,\end{aligned}$$

which is equivalent to

$$\begin{aligned}-(-\alpha+g^Td-v1_m+s)&=(g^T,-1_m,I_m)\begin{pmatrix}d^{\mathrm{new}}\\ v^{\mathrm{new}}\\ s^{\mathrm{new}}\end{pmatrix} &(36)\\ &=(g^T,-1_m)\begin{pmatrix}d^{\mathrm{new}}\\ v^{\mathrm{new}}\end{pmatrix}+s^{\mathrm{new}}\ . &(37)\end{aligned}$$

Linearizing the fourth equation of (23) yields

$$\begin{aligned}0&=F-A_j+d^T\hat{g}_j+\tfrac{1}{2}d^T\hat{G}d+t_j+\nabla_{(d,t_j)}(F-A_j+d^T\hat{g}_j+\tfrac{1}{2}d^T\hat{G}d+t_j)\begin{pmatrix}d^{\mathrm{new}}\\ t_j^{\mathrm{new}}\end{pmatrix}\\ &=F-A_j+d^T\hat{g}_j+\tfrac{1}{2}d^T\hat{G}d+t_j+(\hat{g}_j^T+d^T\hat{G},1)\begin{pmatrix}d^{\mathrm{new}}\\ t_j^{\mathrm{new}}\end{pmatrix}\\ &=F-A_j+d^T\hat{g}_j+\tfrac{1}{2}d^T\hat{G}d+t_j+\big((\hat{g}_j+\hat{G}d)^T,1\big)\begin{pmatrix}d^{\mathrm{new}}\\ t_j^{\mathrm{new}}\end{pmatrix}\quad\text{for all } j\in J_k\\ \iff\ o_m&=F1_m-A+\hat{g}^Td+\tfrac{1}{2}d^T\hat{G}d1_m+t+\big((\hat{g}+\hat{G}d)^T,I_m\big)\begin{pmatrix}d^{\mathrm{new}}\\ t^{\mathrm{new}}\end{pmatrix}\\ &=-A+\hat{g}^Td+(F+\tfrac{1}{2}d^T\hat{G}d)1_m+t+\big((\hat{g}+\hat{G}d)^T,I_m\big)\begin{pmatrix}d^{\mathrm{new}}\\ t^{\mathrm{new}}\end{pmatrix}\ ,\end{aligned}$$

which is equivalent to

$$\begin{aligned}-\big(-A+\hat{g}^Td+(F+\tfrac{1}{2}d^T\hat{G}d)1_m+t\big)&=\big((\hat{g}+\hat{G}d)^T,I_m\big)\begin{pmatrix}d^{\mathrm{new}}\\ t^{\mathrm{new}}\end{pmatrix} &(38)\\ &=(\hat{g}+\hat{G}d)^Td^{\mathrm{new}}+t^{\mathrm{new}}\ . &(39)\end{aligned}$$

Linearizing the fifth equation of (23) yields

$$\begin{aligned}
\sigma c &= \lambda_j s_j + \nabla_{(\lambda_j,s_j)}(\lambda_j s_j)\begin{pmatrix}\lambda_j^{\text{new}}\\ s_j^{\text{new}}\end{pmatrix}\\
&= \lambda_j s_j + (s_j,\lambda_j)\begin{pmatrix}\lambda_j^{\text{new}}\\ s_j^{\text{new}}\end{pmatrix} \quad \text{for all } j\in J_k\\
\iff \quad (s_j,\lambda_j)\begin{pmatrix}\lambda_j^{\text{new}}\\ s_j^{\text{new}}\end{pmatrix} &= -\lambda_j s_j + \sigma c \quad \text{for all } j\in J_k\\
\iff \quad \operatorname{diag}(s)\lambda^{\text{new}} + \operatorname{diag}(\lambda)s^{\text{new}} &= -\operatorname{diag}(\lambda)\operatorname{diag}(s)1_m + \sigma c 1_m\\
\overset{(3)}{\iff} \quad S\lambda^{\text{new}} + \Lambda s^{\text{new}} &= -\Lambda S 1_m + \sigma c 1_m\\
&= -(\Lambda S - \sigma c I_m)1_m && (40)\\
\overset{S^{-1}\cdot}{\iff} \quad I_m\lambda^{\text{new}} + S^{-1}\Lambda s^{\text{new}} &= -\big(\Lambda - \sigma c S^{-1}\big)1_m\ . && (41)
\end{aligned}$$

Linearizing the sixth equation of (23) yields

$$\begin{aligned}
\sigma c &= \mu_j t_j + \nabla_{(\mu_j,t_j)}(\mu_j t_j)\begin{pmatrix}\mu_j^{\text{new}}\\ t_j^{\text{new}}\end{pmatrix}\\
&= \mu_j t_j + (t_j,\mu_j)\begin{pmatrix}\mu_j^{\text{new}}\\ t_j^{\text{new}}\end{pmatrix}\\
\iff \quad (t_j,\mu_j)\begin{pmatrix}\mu_j^{\text{new}}\\ t_j^{\text{new}}\end{pmatrix} &= -\mu_j t_j + \sigma c \quad \text{for all } j\in J_k\\
\iff \quad \operatorname{diag}(t)\mu^{\text{new}} + \operatorname{diag}(\mu)t^{\text{new}} &= -\operatorname{diag}(\mu)\operatorname{diag}(t)1_m + \sigma c 1_m\\
\overset{(3)}{\iff} \quad T\mu^{\text{new}} + M t^{\text{new}} &= -MT 1_m + \sigma c 1_m\\
&= -(MT - \sigma c I_m)1_m && (42)\\
\overset{T^{-1}\cdot}{\iff} \quad I_m\mu^{\text{new}} + T^{-1}M t^{\text{new}} &= -\big(M - \sigma c T^{-1}\big)1_m\ . && (43)
\end{aligned}$$

Now we can show the desired results:

1. Combining (34), (35), (36) and (38) with (40) and (42) yields the general form of the linear system.

2. Combining (34), (35), (36) and (38) with (41) and (43) yields the symmetric form of the linear system.

3. We derive the compact form of the linear system in the following way: Expressing s^{new} in terms of λ^{new} yields

$$\begin{aligned}
S\lambda^{\text{new}} + \Lambda s^{\text{new}} &\overset{(40)}{=} -(\Lambda S - \sigma c I_m)1_m\\
\overset{\Lambda^{-1}\cdot}{\iff} \quad \Lambda^{-1}S\lambda^{\text{new}} + s^{\text{new}} &= -(S - \sigma c\Lambda^{-1})1_m\\
\iff \quad s^{\text{new}} &= -\Lambda^{-1}S\lambda^{\text{new}} - (S - \sigma c\Lambda^{-1})1_m && (44)
\end{aligned}$$

and therefore inserting (44) into (37) yields

$$\begin{aligned}
-(-\alpha + g^T d - v 1_m + s) &\overset{(37)}{=} (g^T, -1_m)\begin{pmatrix}d^{\text{new}}\\ v^{\text{new}}\end{pmatrix} + s^{\text{new}}\\
&\overset{(44)}{=} (g^T, -1_m)\begin{pmatrix}d^{\text{new}}\\ v^{\text{new}}\end{pmatrix} - \Lambda^{-1}S\lambda^{\text{new}} - (S - \sigma c\Lambda^{-1})1_m\\
&= (g^T, -1_m, -\Lambda^{-1}S)\begin{pmatrix}d^{\text{new}}\\ v^{\text{new}}\\ \lambda^{\text{new}}\end{pmatrix} - (S - \sigma c\Lambda^{-1})1_m\ ,
\end{aligned}$$

which is equivalent to

$$
\begin{aligned}
(g^T, -1_m, -\Lambda^{-1}S)\begin{pmatrix} d^{\text{new}} \\ v^{\text{new}} \\ \lambda^{\text{new}} \end{pmatrix} &= -(-\alpha + g^T d - v 1_m + s) + (S - \sigma c \Lambda^{-1}) 1_m \\
&= -\big(-\alpha + g^T d - v 1_m + s - (S - \sigma c \Lambda^{-1}) 1_m \big) \\
&\overset{(3)}{=} -\big(-\alpha + g^T d - v 1_m + \sigma c \Lambda^{-1} 1_m \big) \ . \qquad (45)
\end{aligned}
$$

Furthermore, expressing t^{new} in terms of μ^{new} yields

$$
\begin{aligned}
T\mu^{\text{new}} + Mt^{\text{new}} &\overset{(42)}{=} -(MT - \sigma c I_m) 1_m \\
\overset{M^{-1}\cdot}{\iff} \quad M^{-1}T\mu^{\text{new}} + t^{\text{new}} &= -(T - \sigma c M^{-1}) 1_m \\
\iff \quad t^{\text{new}} &= -M^{-1}T\mu^{\text{new}} - (T - \sigma c M^{-1}) 1_m \qquad (46)
\end{aligned}
$$

and therefore inserting (46) into (39) yields

$$
\begin{aligned}
-\big(-A + \hat{g}^T d + (F + \tfrac{1}{2} d^T \hat{G} d) 1_m + t \big) &\overset{(39)}{=} (\hat{g} + \hat{G}d)^T d^{\text{new}} + t^{\text{new}} \\
&\overset{(46)}{=} (\hat{g} + \hat{G}d)^T d^{\text{new}} - M^{-1}T\mu^{\text{new}} - (T - \sigma c M^{-1}) 1_m \\
&= \big((\hat{g} + \hat{G}d)^T, -M^{-1}T \big) \begin{pmatrix} d^{\text{new}} \\ \mu^{\text{new}} \end{pmatrix} - (T - \sigma c M^{-1}) 1_m \ ,
\end{aligned}
$$

which is equivalent to

$$
\begin{aligned}
\big((\hat{g} + \hat{G}d)^T, -M^{-1}T \big) \begin{pmatrix} d^{\text{new}} \\ \mu^{\text{new}} \end{pmatrix} &= -\big(-A + \hat{g}^T d + (F + \tfrac{1}{2} d^T \hat{G} d) 1_m + t \big) + (T - \sigma c M^{-1}) 1_m \\
&= -\big(-A + \hat{g}^T d + (F + \tfrac{1}{2} d^T \hat{G} d) 1_m + t - (T - \sigma c M^{-1}) 1_m \big) \\
&\overset{(3)}{=} -\big(-A + \hat{g}^T d + (F + \tfrac{1}{2} d^T \hat{G} d) 1_m + \sigma c M^{-1} 1_m \big) \ . \quad (47)
\end{aligned}
$$

4. We derive the practical form of the linear system in the following way: Since

$$
\begin{aligned}
-\big(-\alpha + g^T d - v 1_m + \sigma c \Lambda^{-1} 1_m \big) &\overset{(45)}{=} (g^T, -1_m, -\Lambda^{-1}S) \begin{pmatrix} d^{\text{new}} \\ v^{\text{new}} \\ \lambda^{\text{new}} \end{pmatrix} \\
&= (g^T, -1_m) \begin{pmatrix} d^{\text{new}} \\ v^{\text{new}} \end{pmatrix} - \Lambda^{-1} S \lambda^{\text{new}} \ ,
\end{aligned}
$$

we can express λ^{new} by

$$
\begin{aligned}
\Lambda^{-1}S\lambda^{\text{new}} &= (g^T, -1_m) \begin{pmatrix} d^{\text{new}} \\ v^{\text{new}} \end{pmatrix} - \alpha + g^T d - v 1_m + \sigma c \Lambda^{-1} 1_m \\
\overset{\Lambda S^{-1}\cdot}{\iff} \quad \lambda^{\text{new}} &= \Lambda S^{-1} \Big((g^T, -1_m) \begin{pmatrix} d^{\text{new}} \\ v^{\text{new}} \end{pmatrix} - \alpha + g^T d - v 1_m + \sigma c \Lambda^{-1} 1_m \Big) \\
&= \Lambda S^{-1} (g^T, -1_m) \begin{pmatrix} d^{\text{new}} \\ v^{\text{new}} \end{pmatrix} + \Lambda S^{-1} (-\alpha + g^T d - v 1_m) + \sigma c S^{-1} 1_m \ . \quad (48)
\end{aligned}
$$

Since

$$
\begin{aligned}
-\big(-A + \hat{g}^T d + (F + \tfrac{1}{2} d^T \hat{G} d) 1_m + \sigma c M^{-1} 1_m \big) &\overset{(47)}{=} \big((\hat{g} + \hat{G}d)^T, -M^{-1}T \big) \begin{pmatrix} d^{\text{new}} \\ \mu^{\text{new}} \end{pmatrix} \\
&= (\hat{g} + \hat{G}d)^T d^{\text{new}} - M^{-1}T\mu^{\text{new}} \ ,
\end{aligned}
$$

we can express μ^{new} by

$$
\begin{aligned}
M^{-1}T\mu^{\text{new}} &= (\hat{g} + \hat{G}d)^T d^{\text{new}} - A + \hat{g}^T d + (F + \tfrac{1}{2} d^T \hat{G} d) 1_m + \sigma c M^{-1} 1_m \\
\overset{MT^{-1}\cdot}{\iff} \quad \mu^{\text{new}} &= MT^{-1} \big((\hat{g} + \hat{G}d)^T d^{\text{new}} - A + \hat{g}^T d + (F + \tfrac{1}{2} d^T \hat{G} d) 1_m + \sigma c M^{-1} 1_m \big) \\
&= MT^{-1} (\hat{g} + \hat{G}d)^T d^{\text{new}} + MT^{-1} \big(-A + \hat{g}^T d + (F + \tfrac{1}{2} d^T \hat{G} d) 1_m \big) + \sigma c T^{-1} 1_m \ . \\
&\qquad (49)
\end{aligned}
$$

Now, inserting (48) and (49) into (34) yields

$$
\begin{aligned}
&-\Big(\big(W+\mathrm{tr}(M)\hat G\big)d+g\lambda+\hat g\mu\Big) \overset{(34)}{=} \\
&\overset{(34)}{=} \big(W+\mathrm{tr}(M)\hat G, g, \hat g+\hat G d\big)\begin{pmatrix} d^{\mathrm{new}} \\ \lambda^{\mathrm{new}} \\ \mu^{\mathrm{new}} \end{pmatrix} \\
&= \big(W+\mathrm{tr}(M)\hat G\big)d^{\mathrm{new}} + g\lambda^{\mathrm{new}} + (\hat g+\hat G d)\mu^{\mathrm{new}} \\
&\overset{(48),(49)}{\overset{\downarrow}{=}} \quad \big(W+\mathrm{tr}(M)\hat G\big)d^{\mathrm{new}} \\
&\quad + g\big(\Lambda S^{-1}(g^T,-1_m)\tbinom{d^{\mathrm{new}}}{v^{\mathrm{new}}} + \Lambda S^{-1}(-\alpha+g^T d - v1_m) + \sigma c S^{-1}1_m\big) \\
&\quad + (\hat g+\hat G d)\Big(MT^{-1}(\hat g+\hat G d)^T d^{\mathrm{new}} + MT^{-1}\big(-A+\hat g^T d+(F+\tfrac{1}{2}d^T\hat G d)1_m\big) + \sigma c T^{-1}1_m\Big) \\
&= \big(W+\mathrm{tr}(M)\hat G\big)d^{\mathrm{new}} \\
&\quad + (g\Lambda S^{-1}g^T, -g\Lambda S^{-1}1_m)\tbinom{d^{\mathrm{new}}}{v^{\mathrm{new}}} + g\Lambda S^{-1}(-\alpha+g^T d - v1_m) + \sigma c g S^{-1}1_m \\
&\quad + (\hat g+\hat G d)MT^{-1}(\hat g+\hat G d)^T d^{\mathrm{new}} + (\hat g+\hat G d)MT^{-1}\big(-A+\hat g^T d+(F+\tfrac{1}{2}d^T\hat G d)1_m\big) + \sigma c(\hat g+\hat G d)T \\
&= \Big(\big(W+\mathrm{tr}(M)\hat G\big) + g\Lambda S^{-1}g^T + (\hat g+\hat G d)MT^{-1}(\hat g+\hat G d)^T, -g\Lambda S^{-1}1_m\Big)\tbinom{d^{\mathrm{new}}}{v^{\mathrm{new}}} \\
&\quad + g\Lambda S^{-1}(-\alpha+g^T d - v1_m) + \sigma c g S^{-1}1_m \\
&\quad + (\hat g+\hat G d)MT^{-1}\big(-A+\hat g^T d+(F+\tfrac{1}{2}d^T\hat G d)1_m\big) + \sigma c(\hat g+\hat G d)T^{-1}1_m \ ,
\end{aligned}
$$

which is equivalent to

$$
\begin{aligned}
&\Big(\big(W+\mathrm{tr}(M)\hat G\big) + g\Lambda S^{-1}g^T + (\hat g+\hat G d)MT^{-1}(\hat g+\hat G d)^T, -g\Lambda S^{-1}1_m\Big)\tbinom{d^{\mathrm{new}}}{v^{\mathrm{new}}} = \\
&\quad = -\Big(\big(W+\mathrm{tr}(M)\hat G\big)d+g\lambda+\hat g\mu\Big) \\
&\qquad - \big(g\Lambda S^{-1}(-\alpha+g^T d - v1_m) + \sigma c g S^{-1}1_m\big) \\
&\qquad - \Big((\hat g+\hat G d)MT^{-1}\big(-A+\hat g^T d+(F+\tfrac{1}{2}d^T\hat G d)1_m\big) + \sigma c(\hat g+\hat G d)T^{-1}1_m\Big) \ .
\end{aligned}
\tag{50}
$$

Furthermore, since

$$
\begin{aligned}
1_m^T\Lambda S^{-1}g^T &\overset{(3)}{=} 1_m^T S^{-1}\Lambda g^T \\
&\overset{(3)}{=} 1_m^T S^{-T}\Lambda^T g^T \\
&= (g\Lambda S^{-1}1_m)^T \ ,
\end{aligned}
\tag{51}
$$

we obtain by inserting (48) into (35)

$$
\begin{aligned}
-\big(1-\mathrm{tr}(\Lambda)\big) &\overset{(35)}{=} (-1_m)^T\lambda^{\mathrm{new}} \\
&\overset{(48)}{=} (-1_m)^T\big(\Lambda S^{-1}(g^T,-1_m)\tbinom{d^{\mathrm{new}}}{v^{\mathrm{new}}} + \Lambda S^{-1}(-\alpha+g^T d - v1_m) + \sigma c S^{-1}1_m\big) \\
&= (-1_m^T\Lambda S^{-1}g^T, 1_m^T\Lambda S^{-1}1_m)\tbinom{d^{\mathrm{new}}}{v^{\mathrm{new}}} - 1_m^T\Lambda S^{-1}(-\alpha+g^T d) + v1_m^T\Lambda S^{-1}1_m \\
&\quad - \sigma c 1_m^T S^{-1}1_m \\
&\overset{(22)}{=} \big(-1_m^T\Lambda S^{-1}g^T, \mathrm{tr}(\Lambda S^{-1})\big)\tbinom{d^{\mathrm{new}}}{v^{\mathrm{new}}} - 1_m^T\Lambda S^{-1}(-\alpha+g^T d) + v\cdot\mathrm{tr}(\Lambda S^{-1}) \\
&\quad - \sigma c\cdot\mathrm{tr}(S^{-1}) \\
&\overset{(51)}{=} \big(-(g\Lambda S^{-1}1_m)^T, \mathrm{tr}(\Lambda S^{-1})\big)\tbinom{d^{\mathrm{new}}}{v^{\mathrm{new}}} - 1_m^T\Lambda S^{-1}(-\alpha+g^T d) + v\cdot\mathrm{tr}(\Lambda S^{-1}) \\
&\quad - \sigma c\cdot\mathrm{tr}(S^{-1}) \ ,
\end{aligned}
$$

which is equivalent to

$$\big(-(g\Lambda S^{-1}1_m)^T, \operatorname{tr}(\Lambda S^{-1})\big)\binom{d^{\text{new}}}{v^{\text{new}}} = -\big(1-\operatorname{tr}(\Lambda)\big)+1_m^T\Lambda S^{-1}(-\alpha+g^T d)-v\cdot\operatorname{tr}(\Lambda S^{-1})+\sigma c\cdot\operatorname{tr}(S^{-1})\ . \tag{52}$$

5. We derive the best form of the linear system in the following way: Since

$$\begin{aligned}
&-\big(1-\operatorname{tr}(\Lambda)\big)+1_m^T\Lambda S^{-1}(-\alpha+g^T d)-v\cdot\operatorname{tr}(\Lambda S^{-1})+\sigma c\cdot\operatorname{tr}(S^{-1})\\
&\overset{(52)}{=}\big(-(g\Lambda S^{-1}1_m)^T, \operatorname{tr}(\Lambda S^{-1})\big)\binom{d^{\text{new}}}{v^{\text{new}}}\\
&= -(g\Lambda S^{-1}1_m)^T d^{\text{new}} + \operatorname{tr}(\Lambda S^{-1})v^{\text{new}}
\end{aligned} \tag{53}$$

we can express v^{new} by

$$\operatorname{tr}(\Lambda S^{-1})v^{\text{new}} = (g\Lambda S^{-1}1_m)^T d^{\text{new}} -\big(1-\operatorname{tr}(\Lambda)\big)+1_m^T\Lambda S^{-1}(-\alpha+g^T d)-v\cdot\operatorname{tr}(\Lambda S^{-1})+\sigma c\cdot\operatorname{tr}(S^{-1})$$

which is equivalent to

$$v^{\text{new}} = \frac{(g\Lambda S^{-1}1_m)^T d^{\text{new}}}{\operatorname{tr}(\Lambda S^{-1})} + \frac{-(1-\operatorname{tr}(\Lambda))+1_m^T\Lambda S^{-1}(-\alpha+g^T d)-v\cdot\operatorname{tr}(\Lambda S^{-1})+\sigma c\cdot\operatorname{tr}(S^{-1})}{\operatorname{tr}(\Lambda S^{-1})}\ . \tag{54}$$

Now inserting (54) into (50) yields

$$\begin{aligned}
&-\Big(\big(W+\operatorname{tr}(M)\hat{G}\big)d + g\lambda + \hat{g}\mu\Big)\\
&-\big(g\Lambda S^{-1}(-\alpha+g^T d - v1_m) + \sigma c g S^{-1}1_m\big)\\
&-\Big((\hat{g}+\hat{G}d)MT^{-1}\big(-A+\hat{g}^T d+(F+\tfrac{1}{2}d^T\hat{G}d)1_m\big)+\sigma c(\hat{g}+\hat{G}d)T^{-1}1_m\Big) \overset{(50)}{=}\\
&\overset{(50)}{=}\Big(\big(W+\operatorname{tr}(M)\hat{G}\big)+g\Lambda S^{-1}g^T+(\hat{g}+\hat{G}d)MT^{-1}(\hat{g}+\hat{G}d)^T, -g\Lambda S^{-1}1_m\Big)\binom{d^{\text{new}}}{v^{\text{new}}}\\
&=\Big(\big(W+\operatorname{tr}(M)\hat{G}\big)+g\Lambda S^{-1}g^T+(\hat{g}+\hat{G}d)MT^{-1}(\hat{g}+\hat{G}d)^T\Big)d^{\text{new}} - g\Lambda S^{-1}1_m v^{\text{new}}\\
&\overset{(54)}{=}\Big(\big(W+\operatorname{tr}(M)\hat{G}\big)+g\Lambda S^{-1}g^T+(\hat{g}+\hat{G}d)MT^{-1}(\hat{g}+\hat{G}d)^T\Big)d^{\text{new}}\\
&\quad - g\Lambda S^{-1}1_m\Big(\frac{(g\Lambda S^{-1}1_m)^T d^{\text{new}}}{\operatorname{tr}(\Lambda S^{-1})} + \frac{-(1-\operatorname{tr}(\Lambda))+1_m^T\Lambda S^{-1}(-\alpha+g^T d)-v\cdot\operatorname{tr}(\Lambda S^{-1})+\sigma c\cdot\operatorname{tr}(S^{-1})}{\operatorname{tr}(\Lambda S^{-1})}\\
&=\Big(\big(W+\operatorname{tr}(M)\hat{G}\big)+g\Lambda S^{-1}g^T+(\hat{g}+\hat{G}d)MT^{-1}(\hat{g}+\hat{G}d)^T - \frac{(g\Lambda S^{-1}1_m)(g\Lambda S^{-1}1_m)^T}{\operatorname{tr}(\Lambda S^{-1})}\Big)d^{\text{new}}\\
&\quad - g\Lambda S^{-1}1_m\Big(\frac{-(1-\operatorname{tr}(\Lambda))+1_m^T\Lambda S^{-1}(-\alpha+g^T d)-v\cdot\operatorname{tr}(\Lambda S^{-1})+\sigma c\cdot\operatorname{tr}(S^{-1})}{\operatorname{tr}(\Lambda S^{-1})}\Big)\ ,
\end{aligned}$$

which is equivalent to

$$\begin{aligned}
&\Big(\big(W+\operatorname{tr}(M)\hat{G}\big)+g\Lambda S^{-1}g^T+(\hat{g}+\hat{G}d)MT^{-1}(\hat{g}+\hat{G}d)^T - \frac{(g\Lambda S^{-1}1_m)(g\Lambda S^{-1}1_m)^T}{\operatorname{tr}(\Lambda S^{-1})}\Big)d^{\text{new}}\\
&= -\Big(\big(W+\operatorname{tr}(M)\hat{G}\big)d + g\lambda + \hat{g}\mu\Big)\\
&\quad -\big(g\Lambda S^{-1}(-\alpha+g^T d - v1_m) + \sigma c g S^{-1}1_m\big)\\
&\quad -\Big((\hat{g}+\hat{G}d)MT^{-1}\big(-A+\hat{g}^T d+(F+\tfrac{1}{2}d^T\hat{G}d)1_m\big)+\sigma c(\hat{g}+\hat{G}d)T^{-1}1_m\Big)\\
&\quad + g\Lambda S^{-1}1_m\Big(\frac{-(1-\operatorname{tr}(\Lambda))+1_m^T\Lambda S^{-1}(-\alpha+g^T d)-v\cdot\operatorname{tr}(\Lambda S^{-1})+\sigma c\cdot\operatorname{tr}(S^{-1})}{\operatorname{tr}(\Lambda S^{-1})}\Big)\ .
\end{aligned}$$

6. We derive the best positive definite form of the linear system in the following way: Since

$$
\begin{aligned}
&-\Big(\big(W+\mathrm{tr}(M)\hat{G}\big)d+g\lambda+\hat{g}\mu\Big)\\
&-\big(g\Lambda S^{-1}(-\alpha+g^Td-v1_m)+\sigma cgS^{-1}1_m\big)\\
&-\Big((\hat{g}+\hat{G}d)MT^{-1}\big(-A+\hat{g}^Td+(F+\tfrac{1}{2}d^T\hat{G}d)1_m\big)+\sigma c(\hat{g}+\hat{G}d)T^{-1}1_m\Big)\overset{(50)}{=}\\
&\overset{(50)}{=}\Big(\big(W+\mathrm{tr}(M)\hat{G}\big)+g\Lambda S^{-1}g^T+(\hat{g}+\hat{G}d)MT^{-1}(\hat{g}+\hat{G}d)^T,-g\Lambda S^{-1}1_m\Big)\binom{d^{\text{new}}}{v^{\text{new}}}\\
&=\Big(\big(W+\mathrm{tr}(M)\hat{G}\big)+g\Lambda S^{-1}g^T+(\hat{g}+\hat{G}d)MT^{-1}(\hat{g}+\hat{G}d)^T\Big)d^{\text{new}}-g\Lambda S^{-1}1_mv^{\text{new}}\ .
\end{aligned}
$$

we can express d^{new} by

$$
\begin{aligned}
&-\Big(\big(W+\mathrm{tr}(M)\hat{G}\big)d+g\lambda+\hat{g}\mu\Big)\\
&-\big(g\Lambda S^{-1}(-\alpha+g^Td-v1_m)+\sigma cgS^{-1}1_m\big)\\
&-\Big((\hat{g}+\hat{G}d)MT^{-1}\big(-A+\hat{g}^Td+(F+\tfrac{1}{2}d^T\hat{G}d)1_m\big)+\sigma c(\hat{g}+\hat{G}d)T^{-1}1_m\Big)\\
&+g\Lambda S^{-1}1_mv^{\text{new}}\\
&=\Big(\big(W+\mathrm{tr}(M)\hat{G}\big)+g\Lambda S^{-1}g^T+(\hat{g}+\hat{G}d)MT^{-1}(\hat{g}+\hat{G}d)^T\Big)d^{\text{new}}\ ,
\end{aligned}
$$

which is equivalent to

$$
\begin{aligned}
d^{\text{new}}=&\Big(\big(W+\mathrm{tr}(M)\hat{G}\big)+g\Lambda S^{-1}g^T+(\hat{g}+\hat{G}d)MT^{-1}(\hat{g}+\hat{G}d)^T\Big)^{-1}\cdot\\
&\cdot\Bigg(-\Big(\big(W+\mathrm{tr}(M)\hat{G}\big)d+g\lambda+\hat{g}\mu\Big)\\
&\quad-\big(g\Lambda S^{-1}(-\alpha+g^Td-v1_m)+\sigma cgS^{-1}1_m\big)\\
&\quad-\Big((\hat{g}+\hat{G}d)MT^{-1}\big(-A+\hat{g}^Td+(F+\tfrac{1}{2}d^T\hat{G}d)1_m\big)+\sigma c(\hat{g}+\hat{G}d)T^{-1}1_m\Big)\\
&\quad+g\Lambda S^{-1}1_mv^{\text{new}}\Bigg)\ .
\end{aligned}
\tag{55}
$$

Now inserting (55) into (53) yields

$$
\begin{aligned}
&-\big(1-\operatorname{tr}(\Lambda)\big)+1_m^T\Lambda S^{-1}(-\alpha+g^Td)-v\cdot\operatorname{tr}(\Lambda S^{-1})+\sigma c\cdot\operatorname{tr}(S^{-1})\\
&\overset{(53)}{=}\operatorname{tr}(\Lambda S^{-1})v^{\text{new}}-(g\Lambda S^{-1}1_m)^Td^{\text{new}}\\
&=\operatorname{tr}(\Lambda S^{-1})v^{\text{new}}-(g\Lambda S^{-1}1_m)^T\Big(\big(W+\operatorname{tr}(M)\hat{G}\big)+g\Lambda S^{-1}g^T+(\hat{g}+\hat{G}d)MT^{-1}(\hat{g}+\hat{G}d)^T\Big)^{-1}\cdot\\
&\qquad\cdot\bigg(-\Big(\big(W+\operatorname{tr}(M)\hat{G}\big)d+g\lambda+\hat{g}\mu\Big)\\
&\qquad\quad-\big(g\Lambda S^{-1}(-\alpha+g^Td-v1_m)+\sigma cgS^{-1}1_m\big)\\
&\qquad\quad-\Big((\hat{g}+\hat{G}d)MT^{-1}\big(-A+\hat{g}^Td+(F+\tfrac{1}{2}d^T\hat{G}d)1_m\big)+\sigma c(\hat{g}+\hat{G}d)T^{-1}1_m\Big)\\
&\qquad\quad+g\Lambda S^{-1}1_mv^{\text{new}}\bigg)\\
&=\operatorname{tr}(\Lambda S^{-1})v^{\text{new}}-(g\Lambda S^{-1}1_m)^T\Big(\big(W+\operatorname{tr}(M)\hat{G}\big)+g\Lambda S^{-1}g^T+(\hat{g}+\hat{G}d)MT^{-1}(\hat{g}+\hat{G}d)^T\Big)^{-1}\cdot\\
&\qquad\cdot\bigg(-\Big(\big(W+\operatorname{tr}(M)\hat{G}\big)d+g\lambda+\hat{g}\mu\Big)\\
&\qquad\quad-\big(g\Lambda S^{-1}(-\alpha+g^Td-v1_m)+\sigma cgS^{-1}1_m\big)\\
&\qquad\quad-\Big((\hat{g}+\hat{G}d)MT^{-1}\big(-A+\hat{g}^Td+(F+\tfrac{1}{2}d^T\hat{G}d)1_m\big)+\sigma c(\hat{g}+\hat{G}d)T^{-1}1_m\Big)\bigg)\\
&\qquad-(g\Lambda S^{-1}1_m)^T\Big(\big(W+\operatorname{tr}(M)\hat{G}\big)+g\Lambda S^{-1}g^T+(\hat{g}+\hat{G}d)MT^{-1}(\hat{g}+\hat{G}d)^T\Big)^{-1}g\Lambda S^{-1}1_mv^{\text{new}}\\
&=-(g\Lambda S^{-1}1_m)^T\Big(\big(W+\operatorname{tr}(M)\hat{G}\big)+g\Lambda S^{-1}g^T+(\hat{g}+\hat{G}d)MT^{-1}(\hat{g}+\hat{G}d)^T\Big)^{-1}\cdot\\
&\qquad\cdot\bigg(-\Big(\big(W+\operatorname{tr}(M)\hat{G}\big)d+g\lambda+\hat{g}\mu\Big)\\
&\qquad\quad-\big(g\Lambda S^{-1}(-\alpha+g^Td-v1_m)+\sigma cgS^{-1}1_m\big)\\
&\qquad\quad-\Big((\hat{g}+\hat{G}d)MT^{-1}\big(-A+\hat{g}^Td+(F+\tfrac{1}{2}d^T\hat{G}d)1_m\big)+\sigma c(\hat{g}+\hat{G}d)T^{-1}1_m\Big)\bigg)\\
&\qquad+\bigg(\operatorname{tr}(\Lambda S^{-1})-(g\Lambda S^{-1}1_m)^T\Big(\big(W+\operatorname{tr}(M)\hat{G}\big)+g\Lambda S^{-1}g^T+(\hat{g}+\hat{G}d)MT^{-1}(\hat{g}+\hat{G}d)^T\Big)^{-1}\cdot\\
&\qquad\quad\cdot g\Lambda S^{-1}1_m\bigg)v^{\text{new}}\ ,
\end{aligned}
$$

which is equivalent to

$$v^{\text{new}}=\tfrac{z_1-z_2}{z_3}$$

due to (27), (28), and (29). □

Proposition 2.9 (Search direction for actual implementation)**.**

1. *We set*

$$E := \Lambda S^{-1} \in \mathbb{R}^{m\times m}_{\mathrm{diag}} \tag{56}$$

$$\hat{E} := MT^{-1} \in \mathbb{R}^{m\times m}_{\mathrm{diag}} \tag{57}$$

$$h_{00} := -\alpha + g^T d - v 1_m \in \mathbb{R}^m \text{ (term)} \tag{58}$$

$$h_0 := h_{00} + s \in \mathbb{R}^m \text{ (LS)} \tag{59}$$

$$\hat{\xi}_0 := \hat{G} d \in \mathbb{R}^n \tag{60}$$

$$\hat{h}_{00} := -A + \hat{g}^T d + (F + \tfrac{1}{2} d^T \hat{\xi}_0) 1_m \in \mathbb{R}^m \text{ (term)} \tag{61}$$

$$\hat{h}_0 := \hat{h}_{00} + t \in \mathbb{R}^m \text{ (LS)} \tag{62}$$

$$w_0 := -1 + \mathrm{tr}(\Lambda) \in \mathbb{R} \text{ (LS, term)} \tag{63}$$

and

$$U := W + \mathrm{tr}(M)\hat{G} \in \mathbb{R}^{n\times n}_{\mathrm{sym}} \tag{64}$$

$$u_0 := Ud + (g\lambda + \hat{g}\mu) \in \mathbb{R}^n \text{ (LS, term)} \tag{65}$$

$$h^0_{00} := E(h_{00} + \sigma c \Lambda^{-1} 1_m) \in \mathbb{R}^m \tag{66}$$

$$\hat{h}^0_{00} := \hat{E}(\hat{h}_{00} + \sigma c M^{-1} 1_m) \in \mathbb{R}^m \tag{67}$$

and

$$\hat{\xi}_{00} := \hat{g} + \hat{\xi}_0 \in \mathbb{R}^{n\times m} \ , \tag{68}$$

and we compute

$$B = U + gEg^T + \hat{\xi}_{00}\hat{E}\hat{\xi}_{00}^T \in \mathbb{R}^{n\times n}_{\mathrm{sym}} \text{ positive definite,} \tag{69}$$

and its Cholesky factorization

$$B = LL^T \tag{70}$$

with $L \in \mathbb{R}^{n\times n}_{\mathrm{tril}}$.

2. *We compute*

$$z_1 = w_0 + 1_m^T h^0_{00} \in \mathbb{R} \ ,$$

we set

$$\zeta := gE1_m \in \mathbb{R}^n \tag{71}$$

$$\eta := -u_0 - gh^0_{00} - \hat{\xi}_{00}\hat{h}^0_{00} \in \mathbb{R}^n \ , \tag{72}$$

and we compute

$$-z_2 = (L^{-1}\zeta)^T (L^{-1}\eta) \in \mathbb{R}$$

$$z_3 = \mathrm{tr}(E) - |L^{-1}\zeta|_2^2 \in \mathbb{R}$$

and we set

$$v^{\mathrm{new}} = \frac{z_1 - z_2}{z_3} \in \mathbb{R} \ .$$

3. *We compute*

$$b = \eta + \zeta v^{\mathrm{new}} \in \mathbb{R}^n$$

and we compute d^{new} *by solving the symmetric positive definite linear system*

$$(LL^T) d^{\mathrm{new}} = b \ .$$

4. *We set*

$$h_{11} := g^T d^{\text{new}} - v^{\text{new}} 1_m \in \mathbb{R}^m \tag{73}$$

$$\hat{h}_{11} := \hat{\xi}_{00}^T d^{\text{new}} \in \mathbb{R}^m \tag{74}$$

and we compute

$$\lambda^{\text{new}} = E h_{11} + h_{00}^0 \in \mathbb{R}^m$$
$$\mu^{\text{new}} = \hat{E} \hat{h}_{11} + \hat{h}_{00}^0 \in \mathbb{R}^m \ .$$

5. *We compute*

$$s^{\text{new}} = -E^{-1} \lambda^{\text{new}} - (S - \sigma c \Lambda^{-1}) 1_m \in \mathbb{R}^m$$
$$t^{\text{new}} = -\hat{E}^{-1} \mu^{\text{new}} - (T - \sigma c M^{-1}) 1_m \in \mathbb{R}^m \ .$$

Then $(d^{\text{new}}, v^{\text{new}}, \lambda^{\text{new}}, \mu^{\text{new}}, s^{\text{new}}, t^{\text{new}})$ *solves the best positive definite form of the linear system.*

Proof.

1. We compute

$$\begin{aligned} u_0 &\overset{(65)}{=} Ud + (g\lambda + \hat{g}\mu) \\ &\overset{(64)}{=} \big(W + \operatorname{tr}(M)\hat{G}\big)d + g\lambda + \hat{g}\mu \end{aligned}$$

and

$$\begin{aligned} h_0 &\overset{(59)}{=} h_{00} + s \\ &\overset{(58)}{=} -\alpha + g^T d - v 1_m + s \end{aligned}$$

and

$$\begin{aligned} \hat{h}_0 &\overset{(62)}{=} \hat{h}_{00} + t \\ &\overset{(61)}{=} -A + \hat{g}^T d + (F + \tfrac{1}{2} d^T \hat{\xi}_0) 1_m + t \\ &\overset{(60)}{=} -A + \hat{g}^T d + (F + \tfrac{1}{2} d^T \hat{G} d) 1_m + t \ . \end{aligned}$$

We compute

$$\begin{aligned} B &\overset{(32)}{=} W + \operatorname{tr}(M)\hat{G} + g\Lambda S^{-1} g^T + (\hat{g} + \hat{G}d) M T^{-1} (\hat{g} + \hat{G}d)^T \\ &\overset{(56),(57)}{=} W + \operatorname{tr}(M)\hat{G} + gEg^T + (\hat{g} + \hat{G}d)\hat{E}(\hat{g} + \hat{G}d)^T \\ &\overset{(64)}{=} U + gEg^T + (\hat{g} + \hat{G}d)\hat{E}(\hat{g} + \hat{G}d)^T \\ &\overset{(60)}{=} U + gEg^T + (\hat{g} + \hat{\xi}_0)\hat{E}(\hat{g} + \hat{\xi}_0)^T \\ &\overset{(68)}{=} U + gEg^T + \hat{\xi}_{00} \hat{E} \hat{\xi}_{00}^T \end{aligned}$$

and since $B \in \mathbb{R}^{n \times n}_{\text{sym}}$ from (32) is positive definite, the Cholesky factorization (70) exists.

2. We compute

$$
\begin{aligned}
z_1 &\overset{(28)}{=} -\big(1-\mathrm{tr}(\Lambda)\big) + 1_m^T\Lambda S^{-1}(-\alpha + g^T d) - v\cdot\mathrm{tr}(\Lambda S^{-1}) + \sigma c\cdot\mathrm{tr}(S^{-1})\\
&\overset{(63)}{=} w_0 + 1_m^T\Lambda S^{-1}(-\alpha + g^T d) - v\cdot\mathrm{tr}(\Lambda S^{-1}) + \sigma c\cdot\mathrm{tr}(S^{-1})\\
&\overset{(22)}{=} w_0 + 1_m^T\Lambda S^{-1}(-\alpha + g^T d) - v1_m^T\Lambda S^{-1}1_m + \sigma c\cdot 1_m^T S^{-1}1_m\\
&= w_0 + 1_m^T\Lambda S^{-1}(-\alpha + g^T d - v1_m + \sigma c\cdot\Lambda^{-1}1_m)\\
&\overset{(58)}{=} w_0 + 1_m^T\Lambda S^{-1}(h_{00} + \sigma c\cdot\Lambda^{-1}1_m)\\
&\overset{(56)}{=} w_0 + 1_m^T E(h_{00} + \sigma c\cdot\Lambda^{-1}1_m)\\
&\overset{(66)}{=} w_0 + 1_m^T h_{00}^0
\end{aligned}
$$

and

$$
\begin{aligned}
-z_2 &\overset{(28)}{=} (g\Lambda S^{-1}1_m)^T\big(W + \mathrm{tr}(M)\hat G + g\Lambda S^{-1}g^T + (\hat g + \hat G d)MT^{-1}(\hat g + \hat G d)^T\big)^{-1}\cdot\\
&\quad\cdot\Bigg(-\Big(\big(W + \mathrm{tr}(M)\hat G\big)d + g\lambda + \hat g\mu\Big) - \big(g\Lambda S^{-1}(-\alpha + g^T d - v1_m) + \sigma c g S^{-1}1_m\big)\\
&\qquad - \Big((\hat g + \hat G d)MT^{-1}\big(-A + \hat g^T d + (F + \tfrac12 d^T\hat G d)1_m\big) + \sigma c(\hat g + \hat G d)T^{-1}1_m\Big)\Bigg)\\
&\overset{(64),(56),(57),(60),(68),(69),(65)}{=} (g\Lambda S^{-1}1_m)^T B^{-1}\Bigg(-u_0 - \big(g\Lambda S^{-1}(-\alpha + g^T d - v1_m) + \sigma c g S^{-1}1_m\big)\\
&\qquad - \Big((\hat g + \hat G d)MT^{-1}\big(-A + \hat g^T d + (F + \tfrac12 d^T\hat G d)1_m\big) + \sigma c(\hat g + \hat G d)T^{-1}1_m\Big)\Bigg)\\
&= (g\Lambda S^{-1}1_m)^T B^{-1}\Bigg(-u_0 - g\big(\Lambda S^{-1}(-\alpha + g^T d - v1_m) + \sigma c S^{-1}1_m\big)\\
&\qquad - (\hat g + \hat G d)\Big(MT^{-1}\big(-A + \hat g^T d + (F + \tfrac12 d^T\hat G d)1_m\big) + \sigma c T^{-1}1_m\Big)\Bigg)\\
&= (g\Lambda S^{-1}1_m)^T B^{-1}\Bigg(-u_0 - g\big(\Lambda S^{-1}(-\alpha + g^T d - v1_m + \sigma c\Lambda^{-1}1_m)\big)\\
&\qquad - (\hat g + \hat G d)\Big(MT^{-1}\big(-A + \hat g^T d + (F + \tfrac12 d^T\hat G d)1_m + \sigma c M^{-1}1_m\big)\Big)\Bigg)\\
&\overset{(58),(60),(61)}{=} (g\Lambda S^{-1}1_m)^T B^{-1}\Big(-u_0 - g\big(\Lambda S^{-1}(h_{00} + \sigma c\Lambda^{-1}1_m)\big)\\
&\qquad\qquad - (\hat g + \hat G d)\big(MT^{-1}(\hat h_{00} + \sigma c M^{-1}1_m)\big)\Big)\\
&\overset{(56),(57)}{=} (gE1_m)^T B^{-1}\Big(-u_0 - g\big(E(h_{00} + \sigma c\Lambda^{-1}1_m)\big) - (\hat g + \hat G d)\big(\hat E(\hat h_{00} + \sigma c M^{-1}1_m)\big)\Big)\\
&\overset{(66),(67)}{=} (gE1_m)^T B^{-1}\big(-u_0 - g h_{00}^0 - (\hat g + \hat G d)\hat h_{00}^0\big)\\
&\overset{(70)}{=} (gE1_m)^T (LL^T)^{-1}\big(-u_0 - g h_{00}^0 - (\hat g + \hat G d)\hat h_{00}^0\big)\\
&= (L^{-1}gE1_m)^T L^{-1}\big(-u_0 - g h_{00}^0 - (\hat g + \hat G d)\hat h_{00}^0\big)\\
&\overset{(71)}{=} (L^{-1}\zeta)^T L^{-1}\big(-u_0 - g h_{00}^0 - (\hat g + \hat G d)\hat h_{00}^0\big)\\
&\overset{(60),(68),(72)}{=} (L^{-1}\zeta)^T(L^{-1}\eta)
\end{aligned}
$$

and

$$
\begin{aligned}
z_3 &\overset{(29)}{=} \operatorname{tr}(\Lambda S^{-1}) \\
&\quad - (g\Lambda S^{-1} 1_m)^T \big(W + \operatorname{tr}(M)\hat{G} + g\Lambda S^{-1} g^T + (\hat{g}+\hat{G}d)MT^{-1}(\hat{g}+\hat{G}d)^T\big)^{-1} (g\Lambda S^{-1} 1_m) \\
&\overset{(64),(56),(57),(60),(68),(69)}{\overset{\downarrow}{=}} \quad \operatorname{tr}(E) - (gE1_m)^T B^{-1} (gE1_m) \\
&\overset{(70)}{=} \operatorname{tr}(E) - (gE1_m)^T (LL^T)^{-1} (gE1_m) \\
&= \operatorname{tr}(E) - (L^{-1} gE1_m)^T (L^{-1} gE1_m) \\
&= \operatorname{tr}(E) - |L^{-1} gE1_m|_2^2 \\
&\overset{(71)}{=} \operatorname{tr}(E) - |L^{-1}\zeta|_2^2
\end{aligned}
$$

and now we compute v^{new} according to (30).

3. We compute

$$
\begin{aligned}
b &\overset{(33)}{=} -\Big(\big(W + \operatorname{tr}(M)\hat{G}\big)d + g\lambda + \hat{g}\mu\Big) - \big(g\Lambda S^{-1}(-\alpha + g^T d - v1_m) + \sigma c g S^{-1} 1_m\big) \\
&\quad - \Big((\hat{g}+\hat{G}d)MT^{-1}\big(-A + \hat{g}^T d + (F + \tfrac{1}{2} d^T \hat{G} d)1_m\big) + \sigma c (\hat{g}+\hat{G}d)T^{-1}1_m\Big) \\
&\quad + g\Lambda S^{-1} 1_m v^{\text{new}} \\
&\overset{(64),(65)}{\overset{\downarrow}{=}} \quad -u_0 - g\big(\Lambda S^{-1}(-\alpha + g^T d - v1_m) + \sigma c S^{-1} 1_m - \Lambda S^{-1} 1_m v^{\text{new}}\big) \\
&\quad - (\hat{g}+\hat{G}d)\Big(MT^{-1}\big(-A + \hat{g}^T d + (F + \tfrac{1}{2} d^T \hat{G} d)1_m\big) + \sigma c T^{-1} 1_m\Big) \\
&= -u_0 - g\Big(\Lambda S^{-1}\big((-\alpha + g^T d - v1_m) + \sigma c \Lambda^{-1} 1_m\big) - \Lambda S^{-1} 1_m v^{\text{new}}\Big) \\
&\quad - (\hat{g}+\hat{G}d)\Big(MT^{-1}\big(-A + \hat{g}^T d + (F + \tfrac{1}{2} d^T \hat{G} d)1_m + \sigma c M^{-1} 1_m\big)\Big) \\
&\overset{(58),(60),(61)}{\overset{\downarrow}{=}} \quad -u_0 - g\big(\Lambda S^{-1}(h_{00} + \sigma c \Lambda^{-1} 1_m) - \Lambda S^{-1} 1_m v^{\text{new}}\big) \\
&\qquad - (\hat{g}+\hat{G}d)\big(MT^{-1}(\hat{h}_{00} + \sigma c M^{-1} 1_m)\big) \\
&\overset{(56),(57)}{\overset{\downarrow}{=}} \quad -u_0 - g\big(E(h_{00} + \sigma c \Lambda^{-1} 1_m) - E1_m v^{\text{new}}\big) - (\hat{g}+\hat{G}d)\big(\hat{E}(\hat{h}_{00} + \sigma c M^{-1} 1_m)\big) \\
&\overset{(66),(67)}{\overset{\downarrow}{=}} \quad -u_0 - g(h_{00}^0 - E1_m v^{\text{new}}) - (\hat{g}+\hat{G}d)\hat{h}_{00}^0 \\
&= -u_0 - g h_{00}^0 + gE1_m v^{\text{new}} - (\hat{g}+\hat{G}d)\hat{h}_{00}^0 \\
&\overset{(60),(68),(72)}{\overset{\downarrow}{=}} \quad \eta + gE1_m v^{\text{new}} \\
&\overset{(71)}{=} \eta + \zeta v^{\text{new}}
\end{aligned}
$$

and we solve the linear system

$$
\begin{aligned}
b &\overset{(31)}{=} Bx_{1:n} \\
&\overset{(24)}{=} Bd^{\text{new}} \\
&\overset{(70)}{=} (LL^T)d^{\text{new}} .
\end{aligned}
$$

4. We compute λ^{new} by

$$\begin{aligned}
\lambda^{\text{new}} &\overset{(26)}{=} \Lambda S^{-1}(g^T, -1_m)\binom{d^{\text{new}}}{v^{\text{new}}} + \Lambda S^{-1}(-\alpha + g^T d - v 1_m) + \sigma c S^{-1} 1_m \\
&\overset{(58)}{=} \Lambda S^{-1}(g^T, -1_m)\binom{d^{\text{new}}}{v^{\text{new}}} + \Lambda S^{-1} h_{00} + \sigma c S^{-1} 1_m \\
&= \Lambda S^{-1}(g^T, -1_m)\binom{d^{\text{new}}}{v^{\text{new}}} + \Lambda S^{-1}(h_{00} + \sigma c \Lambda^{-1} 1_m) \\
&\overset{(56)}{=} E(g^T, -1_m)\binom{d^{\text{new}}}{v^{\text{new}}} + E(h_{00} + \sigma c \Lambda^{-1} 1_m) \\
&\overset{(66)}{=} E(g^T, -1_m)\binom{d^{\text{new}}}{v^{\text{new}}} + h_{00}^0 \\
&= E(g^T d^{\text{new}} - 1_m v^{\text{new}}) + h_{00}^0 \\
&\overset{(73)}{=} E h_{11} + h_{00}^0 \ .
\end{aligned}$$

We compute

$$\begin{aligned}
\hat{h}_{11} &\overset{(74)}{=} \hat{\xi}_{00}^T d^{\text{new}} \\
&\overset{(68)}{=} (\hat{g} + \hat{\xi}_0)^T d^{\text{new}} \\
&\overset{(60)}{=} (\hat{g} + \hat{G} d)^T d^{\text{new}}
\end{aligned}$$

and we compute μ^{new} by

$$\begin{aligned}
\mu^{\text{new}} &\overset{(26)}{=} M T^{-1}(\hat{g} + \hat{G} d)^T d^{\text{new}} + M T^{-1}\big(-A + \hat{g}^T d + (F + \tfrac{1}{2} d^T \hat{G} d) 1_m\big) + \sigma c T^{-1} 1_m \\
&\overset{(60),(61)}{=} M T^{-1}(\hat{g} + \hat{G} d)^T d^{\text{new}} + M T^{-1} \hat{h}_{00} + \sigma c T^{-1} 1_m \\
&= M T^{-1}(\hat{g} + \hat{G} d)^T d^{\text{new}} + M T^{-1}(\hat{h}_{00} + \sigma c M^{-1} 1_m) \\
&\overset{(57)}{=} \hat{E}(\hat{g} + \hat{G} d)^T d^{\text{new}} + \hat{E}(\hat{h}_{00} + \sigma c M^{-1} 1_m) \\
&\overset{(67)}{=} \hat{E}(\hat{g} + \hat{G} d)^T d^{\text{new}} + \hat{h}_{00}^0 \\
&\overset{(60),(68),(74)}{=} \hat{E} \hat{h}_{11} + \hat{h}_{00}^0 \ .
\end{aligned}$$

5. We compute s^{new} resp. t^{new} by

$$\begin{aligned}
s^{\text{new}} &\overset{(25)}{=} -\Lambda^{-1} S \lambda^{\text{new}} - (S - \sigma c \Lambda^{-1}) 1_m \\
&\overset{(56)}{=} -E^{-1} \lambda^{\text{new}} - (S - \sigma c \Lambda^{-1}) 1_m
\end{aligned}$$

and

$$\begin{aligned}
t^{\text{new}} &\overset{(25)}{=} -M^{-1} T \mu^{\text{new}} - (T - \sigma c M^{-1}) 1_m \\
&\overset{(57)}{=} -\hat{E}^{-1} \mu^{\text{new}} - (T - \sigma c M^{-1}) 1_m \ .
\end{aligned}$$

□

Proposition 2.10 (Line search scaling & termination criterion for actual implementation)**.** *We assume that the terms of Proposition 2.9 were computed and we additionally set*

$$u_1 := U d^{\text{new}} + \operatorname{tr}(M^{\text{new}}) \hat{\xi}_0 + (g \lambda^{\text{new}} + \hat{g} \mu^{\text{new}}) \in \mathbb{R}^n \text{ (LS)} \tag{75}$$

$$w_1 := \operatorname{tr}(\Lambda^{\text{new}}) \text{ (LS)} \in \mathbb{R} \tag{76}$$

$$h_1 := h_{11} + s^{\text{new}} \in \mathbb{R}^m \text{ (LS)} \tag{77}$$

$$\hat{h}_1 := \hat{h}_{11} + t^{\text{new}} \in \mathbb{R}^m \text{ (LS)} \tag{78}$$

and

$$u_{22} := \hat{G}d^{\mathrm{new}} \in \mathbb{R}^n \tag{79}$$
$$u_2 := \mathrm{tr}(M^{\mathrm{new}})u_{22} \in \mathbb{R}^n \text{ (LS)} \tag{80}$$
$$\hat{h}_2 := (\tfrac{1}{2}d^{\mathrm{new}^T}u_{22})1_m \in \mathbb{R}^m \text{ (LS)} \tag{81}$$

and

$$\varphi_{00} := \Lambda S 1_m \in \mathbb{R}^m \text{ (term)} \tag{82}$$
$$\varphi_0 := \varphi_{00} - \gamma c 1_m \in \mathbb{R}^m \text{ (LS)} \tag{83}$$
$$\hat{\varphi}_{00} := MT1_m \in \mathbb{R}^m \text{ (term)} \tag{84}$$
$$\hat{\varphi}_0 := \hat{\varphi}_{00} - \gamma c 1_m \in \mathbb{R}^m \text{ (LS)} \tag{85}$$
$$\varphi_1 := (\Lambda^{\mathrm{new}}S + \Lambda S^{\mathrm{new}})1_m \in \mathbb{R}^m \text{ (LS)} \tag{86}$$
$$\hat{\varphi}_1 := (M^{\mathrm{new}}T + MT^{\mathrm{new}})1_m \in \mathbb{R}^m \text{ (LS)} \tag{87}$$
$$\chi_1 := 1_m^T\varphi_1 \in \mathbb{R} \tag{88}$$
$$\hat{\chi}_1 := 1_m^T\hat{\varphi}_1 \in \mathbb{R} \tag{89}$$
$$\varphi_2 := \Lambda^{\mathrm{new}}S^{\mathrm{new}}1_m \in \mathbb{R}^m \text{ (LS)} \tag{90}$$
$$\hat{\varphi}_2 := M^{\mathrm{new}}T^{\mathrm{new}}1_m \in \mathbb{R}^m \text{ (LS)} \tag{91}$$
$$\chi_2 := 1_m^T\varphi_2 \in \mathbb{R} \tag{92}$$
$$\hat{\chi}_2 := 1_m^T\hat{\varphi}_2 \in \mathbb{R} \tag{93}$$

and

$$\psi_{00} := \frac{\chi_1 + \hat{\chi}_1}{2m} \in \mathbb{R} \text{ (c-update)} \tag{94}$$
$$\psi_0 := 0.01c + \psi_{00} \in \mathbb{R} \text{ (LS)} \tag{95}$$
$$\psi_1 := \frac{\chi_2 + \hat{\chi}_2}{2m} \in \mathbb{R} \text{ (LS)} \ . \tag{96}$$

1. *For $r \in \mathbb{R}_{\geq 0}$ the the neighbourhood condtion condition for the line search scaling from Proposition 2.4 reads*

$$\left|u_0 + ru_1 + r^2u_2\right|_\infty \leq \beta c$$
$$|w_0 + rw_1| \leq \beta c$$
$$|h_0 + rh_1|_\infty \leq \beta c$$
$$\left|\hat{h}_0 + r\hat{h}_1 + r^2\hat{h}_2\right|_\infty \leq \beta c$$

   and

$$(-r)\varphi_1 - r^2\varphi_2 \leq \varphi_0$$
$$(-r)\hat{\varphi}_1 - r^2\hat{\varphi}_2 \leq \hat{\varphi}_0$$

   and

$$(-r)(\lambda^{\mathrm{new}}, s^{\mathrm{new}}) < (\lambda, s)$$
$$(-r)(\mu^{\mathrm{new}}, t^{\mathrm{new}}) < (\mu, t) \ .$$

2. *For $r \in \mathbb{R}_{>0}$ the centrality reduction condtion for the line search scaling from Proposition 2.4 reads*

$$(-r)\psi_1 \geq \psi_0 \ .$$

and the the centrality parameter c at the next iteration is given by

$$\frac{(\lambda + r\lambda^{\text{new}})^T(s + rs^{\text{new}}) + (\mu + r\mu^{\text{new}})^T(t + rt^{\text{new}})}{2m} = c + r\psi_{00} + r^2\psi_1 \in \mathbb{R} \ .$$

3. *$(d + rd^{\text{new}}, v + rd^{\text{new}}, \lambda + r\lambda^{\text{new}}, \mu + r\mu^{\text{new}}) \in \mathbb{R}^{n+1+2m}$ with $r \in \mathbb{R}_{\geq 0}$ is a primal-dual solution of the reduced QCQP (1) if and only if*

$$\begin{aligned} \left|u_0 + ru_1 + r^2 u_2\right|_\infty &= 0 \\ |w_0 + rw_1| &= 0 \\ |h_0 + rh_1|_\infty &= 0 \\ \left|\hat{h}_0 + r\hat{h}_1 + r^2\hat{h}_2\right|_\infty &= 0 \end{aligned}$$

and

$$\begin{aligned} \varphi_{00} + r\varphi_1 + r^2\varphi_2 &= 0 \\ \hat{\varphi}_{00} + r\hat{\varphi}_1 + r^2\hat{\varphi}_2 &= 0 \end{aligned}$$

and

$$\begin{aligned} (-r)(\lambda^{\text{new}}, s^{\text{new}}) &< (\lambda, s) \\ (-r)(\mu^{\text{new}}, t^{\text{new}}) &< (\mu, t) \ . \end{aligned}$$

Proof.

1. (a) We compute

$$\begin{aligned} u_1 &\overset{(75)}{=} Ud^{\text{new}} + \text{tr}(M^{\text{new}})\hat{\xi}_0 + (g\lambda^{\text{new}} + \hat{g}\mu^{\text{new}}) \\ &\overset{(64)}{=} \big(W + \text{tr}(M)\hat{G}\big)d^{\text{new}} + \text{tr}(M^{\text{new}})\hat{\xi}_0 + g\lambda^{\text{new}} + \hat{g}\mu^{\text{new}} \\ &\overset{(60)}{=} \big(W + \text{tr}(M)\hat{G}\big)d^{\text{new}} + \text{tr}(M^{\text{new}})\hat{G}d + g\lambda^{\text{new}} + \hat{g}\mu^{\text{new}} \end{aligned}$$

and

$$\begin{aligned} u_2 &\overset{(80)}{=} \text{tr}(M^{\text{new}})u_{22} \\ &\overset{(79)}{=} \text{tr}(M^{\text{new}})\hat{G}d^{\text{new}} \end{aligned}$$

and therefore we obtain

$$\begin{aligned} \beta c &\overset{(8)}{\geq} \bigg|\Big(\big(W + \text{tr}(M)\hat{G}\big)d + (g\lambda + \hat{g}\mu)\Big) + r\Big(\big(W + \text{tr}(M)\hat{G}\big)d^{\text{new}} + \text{tr}(M^{\text{new}})\hat{G}d \\ &\qquad + (g\lambda^{\text{new}} + \hat{g}\mu^{\text{new}})\Big) + r^2\text{tr}(M^{\text{new}})\hat{G}d^{\text{new}}\bigg|_\infty \\ &\overset{(64),(65)}{\overset{\downarrow}{=}} \left|u_0 + ru_1 + r^2u_2\right|_\infty \ . \end{aligned}$$

(b) We obtain

$$\begin{aligned} \beta c &\overset{(9)}{\geq} \big|\big(-1 + \text{tr}(\Lambda)\big) + r \cdot \text{tr}(\Lambda^{\text{new}})\big| \\ &\overset{(63),(76)}{\overset{\downarrow}{=}} \left|w_0 + r \cdot \text{tr}(\Lambda^{\text{new}})\right| \ . \end{aligned}$$

(c) We compute

$$\begin{aligned} h_1 &\overset{(77)}{=} h_{11} + s^{\text{new}} \\ &\overset{(73)}{=} g^T d^{\text{new}} - v^{\text{new}} 1_m + s^{\text{new}} \end{aligned}$$

and therefore we obtain

$$\begin{aligned} \beta c &\overset{(10)}{\geq} |(-\alpha + g^T d - v 1_m + s) + r(g^T d^{\text{new}} - v^{\text{new}} 1_m + s^{\text{new}})|_\infty \\ &\overset{(58),(59)}{\overset{\downarrow}{=}} \quad |h_0 + r h_1|_\infty \ . \end{aligned}$$

(d) We compute

$$\begin{aligned} \hat{h}_1 &\overset{(78)}{=} \hat{h}_{11} + t^{\text{new}} \\ &\overset{(74)}{=} (\hat{g} + \hat{G} d)^T d^{\text{new}} + t^{\text{new}} \ . \end{aligned}$$

and

$$\begin{aligned} \hat{h}_2 &\overset{(81)}{=} (\tfrac{1}{2} d^{\text{new}^T} u_{22}) 1_m \\ &\overset{(79)}{=} (\tfrac{1}{2} d^{\text{new}^T} \hat{G} d^{\text{new}}) 1_m \ . \end{aligned}$$

and therefore we obtain

$$\begin{aligned} \beta c &\overset{(11)}{\geq} \left| \big(-A + \hat{g}^T d + (F + \tfrac{1}{2} d^T \hat{G} d) 1_m + t \big) + r \big((\hat{g} + \hat{G} d)^T d^{\text{new}} + t^{\text{new}} \big) + r^2 (\tfrac{1}{2} d^{\text{new}^T} \hat{G} d^{\text{new}}) 1_m \right| \\ &\overset{(60),(61),(62)}{\overset{\downarrow}{=}} \quad \left| \hat{h}_0 + r \hat{h}_1 + r^2 \hat{h}_2 \right|_\infty \ . \end{aligned}$$

(e) We have for all $j = 1, \ldots, m$

$$\begin{aligned} \varphi_0 &\overset{(83)}{=} \varphi_{00} - \gamma c \\ &\overset{(82)}{=} \lambda_j s_j - \gamma c \\ &\overset{(12)}{\geq} (-r)(\lambda_j^{\text{new}} s_j + \lambda_j s_j^{\text{new}}) - r^2 \lambda_j^{\text{new}} s_j^{\text{new}} \\ &\overset{(86),(90)}{\overset{\downarrow}{=}} \quad (-r) \varphi_1 - r^2 \varphi_2 \end{aligned}$$

and

$$\begin{aligned} \hat{\varphi}_0 &\overset{(85)}{=} \hat{\varphi}_{00} - \gamma c \\ &\overset{(84)}{=} \mu_j t_j - \gamma c \\ &\overset{(13)}{\geq} (-r)(\mu_j^{\text{new}} t_j + \mu_j t_j^{\text{new}}) - r^2 \mu_j^{\text{new}} t_j^{\text{new}} \\ &\overset{(87),(91)}{\overset{\downarrow}{=}} \quad (-r) \hat{\varphi}_1 - r^2 \hat{\varphi}_2 \ . \end{aligned}$$

(f) (14) and (15) are clear.

2. We obtain for the centrality reduction condition

$$0 \overset{(16)}{\geq} \left(0.01c + \frac{\lambda^{\text{new}^T} s + \lambda^T s^{\text{new}} + \mu^{\text{new}^T} t + \mu^T t^{\text{new}}}{2m}\right) + r\frac{\lambda^{\text{new}^T} s^{\text{new}} + \mu^{\text{new}^T} t^{\text{new}}}{2m}$$
$$= \left(0.01c + \frac{1_m^T(\Lambda^{\text{new}} S + \Lambda S^{\text{new}})1_m + 1_m^T(M^{\text{new}} T + M T^{\text{new}})1_m}{2m}\right)$$
$$+ r\frac{1_m^T \Lambda^{\text{new}} S^{\text{new}} 1_m + 1_m^T M^{\text{new}} T^{\text{new}} 1_m}{2m}$$
$$\overset{(86),(87),(90),(91)}{=} \left(0.01c + \frac{1_m^T \varphi_1 + 1_m^T \hat{\varphi}_1}{2m}\right) + r\frac{1_m^T \varphi_2 + 1_m^T \hat{\varphi}_2}{2m}$$
$$\overset{(88),(89),(92),(93)}{=} \left(0.01c + \frac{\chi_1 + \hat{\chi}_1}{2m}\right) + r\frac{\chi_2 + \hat{\chi}_2}{2m}$$
$$\overset{(94)}{=} (0.01c + \psi_{00}) + r\frac{\chi_2 + \hat{\chi}_2}{2m}$$
$$\overset{(95),(96)}{=} \psi_0 + r\psi_1$$

and for the value of c at the next iteration

$$\frac{(\lambda + r\lambda^{\text{new}})^T (s + rs^{\text{new}}) + (\mu + r\mu^{\text{new}})^T (t + rt^{\text{new}})}{2m} =$$
$$= \frac{(\lambda^T s + r\lambda^{\text{new}^T} s + r\lambda^T s^{\text{new}} + r^2 \lambda^{\text{new}^T} s^{\text{new}}) + (\mu^T s + r\mu^{\text{new}^T} t + r\mu^T t^{\text{new}} + r^2 \mu^{\text{new}^T} t^{\text{new}})}{2m}$$
$$= \frac{(\lambda^T s + \mu^T s) + r(\lambda^{\text{new}^T} s + \lambda^T s^{\text{new}} + \mu^{\text{new}^T} t + \mu^T t^{\text{new}}) + r^2(\lambda^{\text{new}^T} s^{\text{new}} + \mu^{\text{new}^T} t^{\text{new}})}{2m}$$
$$\overset{(4)}{=} c + \frac{r(\lambda^{\text{new}^T} s + \lambda^T s^{\text{new}} + \mu^{\text{new}^T} t + \mu^T t^{\text{new}}) + r^2(\lambda^{\text{new}^T} s^{\text{new}} + \mu^{\text{new}^T} t^{\text{new}})}{2m}$$
$$\overset{(86),(87),(90),(91),(88),(89),(92),(93)}{=} c + r\frac{\chi_1 + \hat{\chi}_1}{2m} + r^2 \frac{\chi_2 + \hat{\chi}_2}{2m}$$
$$\overset{(94),(96)}{=} c + r\psi_{00} + r^2 \psi_1 \ .$$

3. This is a direct consequence of the fact that $(d, v) \in \mathbb{R}^{n+1}$ solves the reduced QCQP (1) if and only if there exists $(\lambda, \mu, s, t) \in \mathbb{R}^{4m}$ such that (21) holds, the above results, and since we have for all $j = 1, \dots, m$

$$0 \overset{(21)}{\leq} (\lambda_j + r\lambda_j^{\text{new}})(s_j + rs_j^{\text{new}})$$
$$= \lambda_j s_j + r\lambda_j^{\text{new}} s_j + r\lambda_j s_j^{\text{new}} + r^2 \lambda_j^{\text{new}} s_j^{\text{new}}$$
$$= \lambda_j s_j + r(\lambda_j^{\text{new}} s_j + \lambda_j s_j^{\text{new}}) + r^2 \lambda_j^{\text{new}} s_j^{\text{new}}$$
$$\overset{(82),(86),(90)}{=} \varphi_{00} + r\varphi_1 + r^2 \varphi_2$$
$$\iff (-r)\varphi_1 - r^2 \varphi_2 \leq \varphi_{00}$$

and

$$0 \leq (\mu_j + r\mu_j^{\text{new}})(t_j + rt_j^{\text{new}})$$
$$= \mu_j t_j + r\mu_j^{\text{new}} t_j + r\mu_j t_j^{\text{new}} + r^2 \mu_j^{\text{new}} t_j^{\text{new}}$$
$$= \mu_j t_j + r(\mu_j^{\text{new}} t_j + \mu_j t_j^{\text{new}}) + r^2 \mu_j^{\text{new}} t_j^{\text{new}}$$
$$\overset{(84),(87),(91)}{=} \hat{\varphi}_{00} + r\hat{\varphi}_1 + r^2 \hat{\varphi}_2$$
$$\iff (-r)\hat{\varphi}_1 - r^2 \hat{\varphi}_2 \leq \hat{\varphi}_{00} \ . \qquad \square$$

2.4 Algorithm

Now we state the overall (infeasible path following) interior point algorithm for solving the reduced QCQP (1) (cf. WRIGHT [3, p. 166, Algorithm IPF (for convex programming)]).

Algorithm 2.11 (IPF for the reduced QCQP)**.**
Choose $\gamma \in (0,1)$, $\beta > 0$, $0 < \sigma_{\min} < \sigma_{\max} < 1$, $\chi \in (0,1)$, and $(d_0, \lambda_0, \mu_0, s_0, t_0) \in \mathcal{N}_{-\infty}(\gamma, \beta, c_0)$, where c_0 must be chosen according to (4) (e.g., by using (18) and (19)).
`for` $k = 0, 1, 2, \dots$
  Choose $\sigma_k \in [\sigma_{\min}, \sigma_{\max}]$.
  Set v_k^{new} according to (30).
  Solve the (symmetric positive definite) linear system $\bar{B}_k d_k^{\text{new}} = \bar{b}_k$ from (31).
  Set λ_k^{new} and μ_k^{new} according to (26).
  Set s_k^{new} and t_k^{new} according to (25).
  Choose r_k as the first element in the sequence $\{1, \chi, \chi^2, \dots\}$ such that

$$(d_k + rd_k^{\text{new}}, v_k + rv_k^{\text{new}}, \lambda_k + r\lambda_k^{\text{new}}, \mu_k + r\mu_k^{\text{new}}, s_k + rs_k^{\text{new}}, t_k + rt_k^{\text{new}}) \in \mathcal{N}_{-\infty}(\gamma, \beta, c_k)$$

$$\tilde{c}_k(r) := \frac{(\lambda_k + r\lambda_k^{\text{new}})^T (s_k + rs_k^{\text{new}}) + (\mu_k + r\mu_k^{\text{new}})^T (t_k + rt_k^{\text{new}})}{2m} \leq (1 - 0.01r) c_k \ .$$

  Set$(d_{k+1}, v_{k+1}, \lambda_{k+1}, \mu_{k+1}, s_{k+1}, t_{k+1}, c_{k+1}) =$
  $= (d_k + r_k d_k^{\text{new}}, v_k + r_k v_k^{\text{new}}, \lambda_k + r_k \lambda_k^{\text{new}}, \mu_k + r_k \mu_k^{\text{new}}, s_k + r_k s_k^{\text{new}}, t_k + r_k t_k^{\text{new}}, \tilde{c}_k(r_k))$.
`end`

Remark 2.12 (Actual implementation & centring parameter).

- v_k^{new} from (30), d_k^{new} from (31), λ_k^{new}, μ_k^{new} from (26), and s_k^{new} and t_k^{new} from (25) are computed by using the formulas from Proposition (2.9), while the neighbourhood condition and the centrality reduction condition are checked by using the formulas from Proposition 2.10.
- Adaptive strategies and predictor-corrector strategies are the two most important ways for choosing the centring parameter $\sigma_k \in [\sigma_{\min}, \sigma_{\max}]$ (s. NOCEDAL & WRIGHT [2, p. 572 (pdf: 591)]).

References

[1] H. Fendl. *A feasible second order bundle algorithm for nonsmooth, nonconvex optimization problems with inequality constraints and its application to certificates of infeasibility.* PhD thesis, Universität Wien, 2011.

[2] J. Nocedal and S.J. Wright. *Numerical Optimization.* Springer Series in Operations Research and Financial Engineering. Springer, New York, 2nd ed., 2006.

[3] S.J. Wright. *Primal-dual interior-point methods.* SIAM, Philadelphia, PA, 1997.