Supplementary Material for *Accelerated Training for Matrix-norm Regularization: A Boosting Approach*

A Proof of Theorem 1

In this section we prove the $O(1/\epsilon)$ convergence rate of the boosting Algorithm 1.

Theorem 1 (Rate of convergence) *Under Assumption 1, Algorithm 1 finds an ϵ accurate solution to (5) in $O(1/\epsilon)$ number of steps. More precisely, denoting f^* as the minimum of (5), then*

$$f(\{\sigma_i^{(k)}, A_i^{(k)}\}) - f^* \le \frac{4C_L}{k+2}.$$

Proof: Denoting $s^* = \sum_i \sigma_i^*$, where recall that $\{A_i^*, \sigma_i^*\}$ is some optimal solution to (5). Our proof is based upon the following observation:

$$\begin{aligned} f^* &= \min_{A_i \in \mathcal{A}, \sigma_i \ge 0} L\left(\sum_i \sigma_i A_i\right) + \lambda \sum_i \sigma_i \\ &= \min_{Y \in s^* \mathcal{K}} L(Y) + \lambda s^*, \end{aligned} \tag{19}$$

where $\mathcal{K}$ is the convex hull of the set $\mathcal{A}$.

Let $s_k := \sum_i \sigma_i^{(k)}$. We prove Theorem 1 for a "weaker" version of Algorithm 1, where a_k is set to some constant $1 - \eta_k$. The following chain of inequalities consists the main part of our proof.

$$\begin{aligned}
f(X_k) &= L(X_k) + \lambda s_k \\
\text{(Definition of } X_k, s_k) &= \min_{\rho \ge 0} L\left((1-\eta_k)X_{k-1} + \rho\eta_k H_k\right) + \lambda(1-\eta_k)s_{k-1} + \lambda\rho\eta_k && (20) \\
&\le L((1-\eta_k)X_{k-1} + \eta_k(s^* H_k)) + \lambda(1-\eta_k)s_{k-1} + \lambda s^* \eta_k \\
\text{(Assumption 1)} &\le f(X_{k-1}) + \eta_k \langle s^* H_k - X_{k-1}, \nabla L(X_{k-1})\rangle + \frac{C_L}{2}\eta_k^2 - \lambda\eta_k s_{k-1} + \lambda\eta_k s^* && (21) \\
\text{(Definition of } H_k) &\le \min_{Y \in s^* \cdot \mathcal{A}} f(X_{k-1}) + \eta_k \langle Y - X_{k-1}, \nabla L(X_{k-1})\rangle + \frac{C_L}{2}\eta_k^2 - \lambda\eta_k s_{k-1} + \lambda\eta_k s^* \\
\text{(Linearity)} &\le \min_{Y \in s^* \cdot \mathcal{K}} f(X_{k-1}) + \eta_k \langle Y - X_{k-1}, \nabla L(X_{k-1})\rangle + \frac{C_L}{2}\eta_k^2 - \lambda\eta_k s_{k-1} + \lambda\eta_k s^* && (22) \\
\text{(Convexity of } L) &\le \min_{Y \in s^* \cdot \mathcal{K}} f(X_{k-1}) + \eta_k(L(Y) - L(X_{k-1})) + \frac{C_L}{2}\eta_k^2 - \lambda\eta_k s_{k-1} + \lambda\eta_k s^* \\
\text{(Rearrangement)} &= (1-\eta_k)f(X_{k-1}) + \eta_k \min_{Y \in s^* \cdot \mathcal{K}} \left(L(Y) + \lambda s^*\right) + \frac{C_L}{2}\eta_k^2 \\
\text{(Observation (19))} &= (1-\eta_k)f(X_{k-1}) + \eta_k f^* + \frac{C_L}{2}\eta_k^2,
\end{aligned}$$

hence

$$f(X_k) - f^* \le (1-\eta_k)(f(X_{k-1}) - f^*) + \frac{C_L}{2}\eta_k^2.$$

Setting $\eta_k = \frac{2}{k+2}$, and an easy induction argument establishes that

$$f(X_k) - f^* \le \frac{4C_L}{k+2}.$$

■

The proof, although completely elementary, does harness several interesting ideas. Note first that in, say, the analysis of the ordinary gradient algorithm, one usually upper bounds the convex function L by its quadratic expansion

$$L(Y) \le L(X) + \langle Y - X, \nabla L(X)\rangle + \frac{\hat{C}_L}{2}\|Y - X\|^2,$$

and then tries to minimize the quadratic upper bound; while in contrast, our analysis above takes perhaps a surprisingly loose step: upper bound L by the *linear* function

$$L(y) \le L(x) + \langle Y - X, \nabla L(X)\rangle + \frac{C_L}{2}.$$

The (huge) gain, of course, is the possibility of inequality (22), which allows us to select the next update by optimizing over the (potentially much simpler) set $\mathcal{A}$, instead of the convex hull $\mathcal{K}$.

The next key ingredient in the proof is our observation (19), which is completely trivial, yet after combining it with the one dimensional line search over $\rho \ge 0$ (or b in Algorithm 1), Algorithm 1 behaves as if it knew the *unknown* but fixed constant s^*.

Some remarks regarding to Theorem 1 are in order.

- If the loss function L is only Lipchitz continuous, then one can apply the "smoothing" trick [35] to get $O(\frac{1}{\epsilon^2})$ convergence rate for algorithm 1.
- Our result heavily builds on previous work [10, 18], however, it seems that our treatment is slightly more general. For instance, the ℓ_1 norm regularizer $\sum_i \sigma_i$ can be readily replaced by $h(\sum_i \sigma_i)$, where $h : \mathbb{R}_+ \to \mathbb{R}$ is some convex function. Essentially the same proof would still go through. Take h as the indicator of some convex set recovers most previous results, which all consider the constrained problem instead of the arguably more natural regularized problem[6].
- The line search step in Algorithm 1 need not be solved exactly. We can derive essentially the same rate as long as the error decays at the rate $O(\frac{1}{k})$.
- The step size $\eta_k = O(\frac{1}{k})$ is optimal, among constant ones, in the following sense. We usually prefer large step sizes since they often than not result in faster convergence; on the other hand, Algorithm 1 needs to be able to reset any σ_i to 0, which requires that the discount factor $\prod_{k=1}^{\infty}(1-\eta_k) = 0$. It is not hard to show that the latter condition is satisfied iff $\sum_{k=1}^{\infty} \eta_k = \infty$, hence the near optimality of the step size $O(\frac{1}{k})$.

B Improved Rate When L is Strongly Convex

In this section, under an additional assumption, we improve the convergence rate in Theorem 1 by considering the totally corrective algorithm in (7).

Recall that strong convexity (with modulus μ) of L implies that

$$L(Y) \ge L(X) + \langle Y - X, \nabla L(X)\rangle + \frac{\mu}{2}\|L - X\|^2. \tag{23}$$

Note that the constant μ depends on the choice of the norm $\|\cdot\|$. In the proof we fix the norm to be essentially ℓ_1, and we assume the set $\mathcal{A}$ consists of finitely many points.

Theorem 2 *Suppose Assumption 1 holds and L is furthermore strongly convex with modulus μ. Let $\{A_i^*, \sigma_i^*\}$ be a minimzer of (5) and denote $f^* := f(\{A_i^*, \sigma_i^*\})$, $s^* := \sum_i \sigma_i^*$. Then the totally corrective algorithm converges at least linearly. More precisely*

$$f(\{\sigma_i^{(k)}, A_i^{(k)}\}) - f^* \le \left(1 - \min\left\{\frac{1}{2}, \frac{2\mu(s^*)^2}{m^2 C_L}\right\}\right)^k (f(\{0, \mathbf{0}\}) - f^*),$$

where m is the number of non-zeros in $\{\sigma_i^*\}$.

Our proof is essentially in the same spirit as that of [16, Theorem 2.8], see also [17, Theorem 2]. It is a pleasant surprise that the latter proof extends without much difficulty to the regularized problem considered here.

Proof: In the proof we will use $f(X_k)$ to denote $L(X_k) + \lambda \sum_i \sigma_i^{(k)}$ where $X_k := \sum_{i=1}^k \sigma_i^{(k)} A_i^{(k)}$.

[6] After completion of the first draft, we became aware of the recent paper [15], which proposed an algorithm similar as our totally corrective version in (7) for the regularized problem, but the rate proven there, $O(\frac{1}{\epsilon^2})$, is worse than the one presented in our Theorem 1.

Let us record the optimality condition in (7): $\forall\, \boldsymbol{\tau} \in \mathbb{R}_+^k$, the following holds

$$\sum_{i=1}^{k} (\langle A_i^{(k)}, \nabla L(X_k)\rangle + \lambda)(\tau_i - \sigma_i^{(k)}) \geq 0, \tag{24}$$

where $\sigma_i^{(k)}$ denotes the optimal solution in (7).

Take $0 \leq \eta \leq 1$ whose value will be optimized later. Let $s_k := \sum_{i=1}^k \sigma_i^{(k)}$. From Assumption 1 we have

$$\begin{aligned} f((1-\eta)X_k + \eta s^* H_{k+1}) &= L((1-\eta)X_k + \eta s^* H_{k+1}) + (1-\eta)\lambda s_k + \eta\lambda s^* \\ &\leq f(X_k) + \eta \langle s^* H_{k+1} - X_k, \nabla L(X_k)\rangle + \frac{C_L}{2}\eta^2 + \eta\lambda(s^* - s_k). \end{aligned} \tag{25}$$

We need to define two index sets I and J, where I contains the indexes of the elements in $\{A_i^*\}$ but not in $\{A_i^{(k)}\}$ while J contains the indexes of the elements in both $\{A_i^*\}$ and $\{A_i^{(k)}\}$. Note that we can assume that I is nonempty since otherwise the current totally corrective step will find an optimal solution.

Define $r = \sum_{i \in I} \sigma_i^*$, and by the definition of H_{k+1},

$$\begin{aligned} r\langle s^* H_{k+1}, \nabla L(X_k)\rangle &\leq \sum_{i\in I} s^* \sigma_i^* \langle A_i, \nabla L(X_k)\rangle \\ &= \sum_{i\in I} (s^*\sigma_i^* - (s^* - r)\sigma_i^{(k)}) \langle A_i, \nabla L(X_k)\rangle \\ &\leq \sum_{i\in J} (s^*\sigma_i^* - (s^* - r)\sigma_i^{(k)}) \langle A_i, \nabla L(X_k)\rangle + \lambda(s^* - r)(s^* - s_k) \\ &= s^*(\langle X^* - X_k, \nabla L(X_k)\rangle + \lambda(s^* - s_k)) - \lambda r(s^* - s_k) + r\langle X_k, \nabla L(X_k)\rangle \\ &\leq s^*(f^* - f(X_k) - \frac{\mu}{2}\|\sigma^* - \sigma^{(k)}\|_1^2) - \lambda r(s^* - s_k) + r\langle X_k, \nabla L(X_k)\rangle, \end{aligned} \tag{26}$$

where the last inequality follows from the strong convexity assumption, and the second inequality follows from the optimality of $\sigma^{(k)}$. Indeed, if $J - I = \emptyset$, then $s^* = r$, hence we in fact have an equality. Assume otherwise, then the inequality follows from the optimality condition (24).

Now apply (26) to (25), we get

$$\begin{aligned} f((1-\eta)X_k + \eta s^* H_{k+1}) &\leq f(X_k) + \eta \frac{r\langle s^* H_{k+1}, \nabla L(X_k)\rangle - r\langle X_k, \nabla L(X_k)\rangle}{r} + \frac{C_L}{2}\eta^2 + \eta\lambda(s^* - s_k) \\ &\leq f(X_k) - \eta \frac{s^*(f(X_k) - f^* + \frac{\mu}{2}\|\sigma^* - \sigma^{(k)}\|_1^2)}{r} + \frac{C_L}{2}\eta^2. \end{aligned}$$

Apparently $f(X_{k+1}) \leq \min_{\eta\in[0,1]} f((1-\eta)X_k + \eta s^* H_{k+1})$, hence

$$f(X_{k+1}) - f^* \leq f(X_k) - f^* - \eta \frac{s^*(f(X_k) - f^* + \frac{\mu}{2}\|\sigma^* - \sigma^{(k)}\|_1^2)}{r} + \frac{C_L}{2}\eta^2.$$

Minimizing η on the right-hand side yields

$$f(X_{k+1}) - f^* \leq f(X_k) - f^* - \min\left\{\frac{s^*\delta}{2r}, \frac{\delta^2 (s^*)^2}{2r^2 C_L}\right\},$$

where $\delta := f(X_k) - f^* + \frac{\mu}{2}\|\sigma^* - \sigma^{(k)}\|_1^2 \geq 0$. It is easy to see that

$$\frac{s^*\delta}{2r} \geq \frac{1}{2}(f(X_k) - f^*).$$

On the other hand,

$$\begin{aligned}\frac{\delta^2(s^*)^2}{2r^2C_L} &\geq \frac{2\mu(f(X_k)-f^*)(s^*)^2\|\sigma^*-\sigma^{(k)}\|_1^2}{r^2C_L} \geq \frac{2\mu(f(X_k)-f^*)(s^*)^2\sum_{i\in I}(\sigma_i^*)^2}{C_L(\sum_{i\in I}\sigma_i^*)^2} \\ &\geq \frac{2\mu(f(X_k)-f^*)(s^*)^2}{C_L|I|^2} \geq \frac{2\mu(f(X_k)-f^*)(s^*)^2}{C_Lm^2} = \frac{2\mu(s^*)^2}{C_Lm^2}(f(X_k)-f^*),\end{aligned} \tag{27}$$

where recall that m is the number of nonzeros entries in $\{\sigma_i^*\}$.

Combining the above two estimates completes the proof:

$$f(X_{k+1})-f^* \leq \left(1-\min\left\{\frac{1}{2}, \frac{2\mu(s^*)^2}{C_Lm^2}\right\}\right)(f(X_k)-f^*).$$

■

C Proof of Proposition 1 and 2

Recall that $\mathcal{K}$ is the convex hull of $\mathcal{A}$.

Proposition 1 $\gamma_{\mathcal{K}}(X) = \min\limits_{U,V:UV=X} \frac{1}{2}\sum_i(\|U_{:i}\|_C^2+\|V_{i:}\|_R^2) = \min\limits_{U,V:UV=X}\sum_i \|U_{:i}\|_C\,\|V_{i:}\|_R$.

Proof: This proof is similar in spirit to [36]. For any $UV=X$, we can write

$$X = \sum_i \|U_{:i}\|_C\,\|V_{i:}\|_R\,\frac{U_{:i}}{\|U_{:i}\|_C}\frac{V_{i:}}{\|V_{i:}\|_R}. \tag{28}$$

So by the definition of gauge function,

$$\gamma_{\mathcal{K}}(X) \leq \sum_i \|U_{:i}\|_C\,\|V_{i:}\|_R \leq \frac{1}{2}\sum_i\left(\|U_{:i}\|_C^2+\|V_{i:}\|_R^2\right). \tag{29}$$

To attain equality, by the the definition of the gauge $\gamma_{\mathcal{K}}$, there exist σ_i, $\hat{U}$, and $\hat{V}$ which satisfy

$$\|\hat{U}_{:i}\|_C = \|\hat{V}_{i:}\|_R = 1, \qquad \sum_i \sigma_i\hat{U}_{:i}\hat{V}_{i:} = X, \qquad \gamma_{\mathcal{K}}(X) = \sum_i\sigma_i, \qquad \sigma_i \geq 0. \tag{30}$$

Then define $U_{:i} = \sqrt{\sigma_i}\hat{U}_{:i}$ and $V_{i:} = \sqrt{\sigma_i}\hat{V}_{i:}$. It is easy to verify that $UV = X$ and $\frac{1}{2}(\|U_{:i}\|_C^2+\|V_{i:}\|_R^2) = \sum_i\|U_{:i}\|_C\,\|V_{i:}\|_R = \sum_i\sigma_i = \gamma_{\mathcal{K}}(X)$. ■

Proposition 2 *For any $U\in\mathbb{R}^{m\times k}, V\in\mathbb{R}^{k\times n}$, there exist $\alpha_i\geq 0$, $\|\boldsymbol{\alpha}\|_0\leq k$ and $\mathbf{u}_i$, $\mathbf{v}_i$ such that*

$$UV = \sum_i\alpha_i\mathbf{u}_i\mathbf{v}_i', \quad \|\mathbf{u}_i\|_C\leq 1, \quad \|\mathbf{v}_i\|_R\leq 1, \quad \sum_i\alpha_i = \frac{1}{2}\sum_i(\|U_{:i}\|_C^2+\|V_{i:}\|_R^2).$$

Proof: Denote $a_i = \|U_{:i}\|_C$ and $b_i = \|V_{i:}\|_R$. Then

$$UV = \sum_i a_ib_i\frac{U_{:i}}{a_i}\frac{V_{i:}}{b_i} = \sum_i \underbrace{\frac{1}{2}(a_i^2+b_i^2)}_{:=\alpha_i}\underbrace{\sqrt{\frac{a_ib_i}{\frac{1}{2}(a_i^2+b_i^2)}}\frac{U_{:i}}{a_i}}_{:=\mathbf{u}_i}\underbrace{\sqrt{\frac{a_ib_i}{\frac{1}{2}(a_i^2+b_i^2)}}\frac{V_{i:}}{b_i}}_{:=\mathbf{v}_i'}. \tag{31}$$

Clearly $\|\mathbf{u}_i\|_C\leq 1$, $\|\mathbf{v}_i\|_R\leq 1$, and $\sum_i\alpha_i = \frac{1}{2}\sum_i(\|U_{:i}\|_C^2+\|V_{i:}\|_R^2)$. ■

D Proof of the strong duality

The goal of this note is to solve the following problem:

$$(QP)\ \max_{\mathbf{x},\mathbf{y}} \|A\mathbf{x}+B\mathbf{y}+\mathbf{c}\|\,,\ s.t.\ \|\mathbf{x}\|\le 1,\ \|\mathbf{y}\|\le 1. \tag{32}$$

Here $\mathbf{c}$ is a non-zero vector, and all the norms are Euclidean. Let $\mathbf{x}\in\mathbb{R}^m$, $\mathbf{y}\in\mathbb{R}^n$, $\mathbf{c}\in\mathbb{R}^t$, $A\in\mathbb{R}^{t\times m}$ and $B\in\mathbb{R}^{t\times n}$.

The problem is not convex in this form, because it is *maximizing* a positive semi-definite quadratic. To find a global solution, we first reformulate it. Define

$$\mathbf{z}=\begin{pmatrix} r\\ \mathbf{x}\\ \mathbf{y}\end{pmatrix},\quad \mathbf{b}=\begin{pmatrix} A'\mathbf{c}\\ B'\mathbf{c}\end{pmatrix} \tag{33}$$

$$Q=-\begin{pmatrix} A'A & A'B\\ B'A & B'B\end{pmatrix},\ M_0=\begin{pmatrix} 0 & -\mathbf{b}'\\ -\mathbf{b} & Q\end{pmatrix} \tag{34}$$

$$M_1=\begin{pmatrix} -1 & \mathbf{0}_{1\times m} & \mathbf{0}_{1\times n}\\ \mathbf{0}_{m\times 1} & I_{m\times m} & \mathbf{0}_{m\times n}\\ \mathbf{0}_{n\times 1} & \mathbf{0}_{n\times m} & \mathbf{0}_{n\times n}\end{pmatrix} \tag{35}$$

$$M_2=\begin{pmatrix} -1 & \mathbf{0}_{1\times m} & \mathbf{0}_{1\times n}\\ \mathbf{0}_{m\times 1} & \mathbf{0}_{m\times m} & \mathbf{0}_{m\times n}\\ \mathbf{0}_{n\times 1} & \mathbf{0}_{n\times m} & I_{n\times n}\end{pmatrix}. \tag{36}$$

Then the problem (QP) can be rewritten as

$$(QP)\quad \max_{\mathbf{z}}\quad \mathbf{z}'M_0\mathbf{z} \tag{37}$$

$$s.t.\quad \mathbf{z}'M_1\mathbf{z}\le 0 \tag{38}$$

$$\mathbf{z}'M_2\mathbf{z}\le 0 \tag{39}$$

$$r^2=1. \tag{40}$$

Denote the inner product between matrices X and Y as $X\bullet Y := \operatorname{tr} X'Y$. Then we can further rewrite (QP) as:

$$(QP)\quad \min_{\mathbf{z}}\quad M_0\bullet(\mathbf{z}\mathbf{z}') \tag{41}$$

$$s.t.\quad M_1\bullet(\mathbf{z}\mathbf{z}')\le 0 \tag{42}$$

$$M_2\bullet(\mathbf{z}\mathbf{z}')\le 0 \tag{43}$$

$$I_{00}\bullet(\mathbf{z}\mathbf{z}')=1, \tag{44}$$

where $I_{00}=\begin{pmatrix} 1 & \mathbf{0}_{1\times(m+n)}\\ \mathbf{0}_{(m+n)\times 1} & \mathbf{0}_{(m+n)\times(m+n)}\end{pmatrix}$. Then a natural SDP relaxation of (QP) is

$$(SP)\quad \min_{X}\quad M_0\bullet X \tag{45}$$

$$s.t.\quad M_1\bullet X\le 0 \tag{46}$$

$$M_2\bullet X\le 0 \tag{47}$$

$$I_{00}\bullet X=1, \tag{48}$$

$$X\succeq\mathbf{0}. \tag{49}$$

Note (SP) is a convex problem, but there may be a gap between the optimal values of (SP) and (QP) because (SP) dropped the rank-one constraint on X. The dual problem of (SP) is

$$(SD)\quad \max_{y_0,y_1,y_2}\quad y_0 \tag{50}$$

$$s.t.\quad Z:=M_0-y_0I_{00}+y_1M_1+y_2M_2\succeq\mathbf{0} \tag{51}$$

$$y_1\ge 0,\ y_2\ge 0. \tag{52}$$

With slight abuse of notation, we denote as QP, SP, and SD the optimal objective value of the respective problems. We may also write $QP(A,B,\mathbf{c})$ to make explicit their dependence on $(A,B,\mathbf{c})$.

Clearly $SP = SD$ since the Slater's condition is always met. However, $QP \geq SP$ because (SP) does not necessarily admit a rank-one optimal solution. The key conclusion of this note is to rule out this possibility, and show that

$$QP(A, B, \mathbf{c}) = SP(A, B, \mathbf{c}) \quad \text{for all } A, B, \mathbf{c}, \quad \text{(strong duality).} \tag{53}$$

So there must be a rank-one optimal solution to (SP), based on which we can easily recover an optimal $\mathbf{z}$ for (QP).

Generalization Note the $\mathbf{b}$ in (33) is determined by A, B and $\mathbf{c}$ and does not have full freedom. In this note we will prove a stronger result by dropping this constraint and consider for general unconstrained $\mathbf{b}$. Accordingly, we will show a slightly more general relationship:

$$QP(A, B, \mathbf{b}) = SP(A, B, \mathbf{b}) \quad \text{for all } A, B, \mathbf{b}, \quad \text{(strong duality).} \tag{54}$$

Besides proving (54), two computational issues need to be resolved. First, given the optimal $\{y_i\}$ for (SD), how to recover the optimal $(\mathbf{x}, \mathbf{y})$ for (QP). The details are given in Section D.1.3. Second, how to solve (SD). We propose using the cutting plane method. Note there are only three variables in (SD), and the only tricky part is the positive semi-definite constraint (51). For low dimensional convex optimization, it is quite easy to approximate this (nontrivial) constraint by cutting planes, which relies on the oracle: given an assignment of $\{y_i\}$, find a maximal violator of (51), *i.e.* $\text{argmin}_{\mathbf{u}:\|\mathbf{u}\|=1} \mathbf{u}' Z \mathbf{u}$ (≤ 0). The solution is simply the eigenvector corresponding to the least algebraically eigenvalue.

Notation The set of all n-by-n symmetric matrices is denoted as $\mathcal{S}^{n\times n}$, and the set of all n-by-n positive semi-definite matrices is denoted as $\mathcal{S}_+^{n\times n}$. $\det(A)$ is the determinant of a matrix A. Denote the kernel (null space) of a linear map A as $\text{Ker}(A)$, and the range of A as $\text{Im}(A)$ (the span of the column space of A).

D.1 Strong Duality

This section proves the strong duality. Our idea is similar to [37]. We first define a set of Properties (called Property $\mathcal{I}$) over the optimal solutions of (SP) and (SD). Next we show that if Property $\mathcal{I}$ does not hold, then strong duality is guaranteed. Finally we show that in our case, Property $\mathcal{I}$ can never be met.

D.1.1 Property $\mathcal{I}$

Let $\hat{X}$ and $(\hat{Z}, \hat{y}_0, \hat{y}_1, \hat{y}_2)$ be a pair of optimal solutions for (SP) and (SD), respectively. The KKT condition states

$$\hat{X}\hat{Z} = \mathbf{0} \tag{55}$$

$$\hat{y}_i M_i \bullet \hat{X} = 0, \qquad i \in \{1, 2\}. \tag{56}$$

We define a Property $\mathcal{I}$ in the same spirit as [37].

Definition 1 *We say the optimal pair $\hat{X}$ and $(\hat{Z}, \hat{y}_0, \hat{y}_1, \hat{y}_2)$ has Property $\mathcal{I}$ if:*

1. $M_1 \bullet \hat{X} = 0$ *and* $M_2 \bullet \hat{X} = 0$.
2. $\text{rank}(\hat{Z}) = m + n - 1$.
3. $\text{rank}(\hat{X}) = 2$, *and P3: there is a rank-one decomposition of $\hat{X}$, $\hat{X} = \mathbf{x}_1\mathbf{x}_1' + \mathbf{x}_2\mathbf{x}_2'$, such that $M_1 \bullet \mathbf{x}_i\mathbf{x}_i' = 0$ $(i = 1, 2)$, and $(M_2 \bullet \mathbf{x}_1\mathbf{x}_1')(M_2 \bullet \mathbf{x}_2\mathbf{x}_2') < 0$.*

The concept of rank-one decomposition is available in subsection D.2. It is simple to symmetrize the item 3 of Property $\mathcal{I}$ (*i.e.* swap the role of M_1 and M_2), but this is not needed for our purposes. Our key result is to use the Property $\mathcal{I}$ to characterize the case of strong duality.

Theorem 3 *If (SP) and (SD) have a pair of optimal solution $\hat{X}$ and $(\hat{Z}, \hat{y}_0, \hat{y}_1, \hat{y}_2)$ which do not satisfy Property $\mathcal{I}$, then strong duality holds,* i.e. *$SP = QP$ and (SP) has a rank-one optimal solution.*

Proof: Assume Property $\mathcal{I}$ does not hold, and we enumerate four exhaustive (but not mutually exclusive) possibilities.

Case 1: $M_1 \bullet \hat{X} \neq 0$ or $M_2 \bullet \hat{X} \neq 0$. Without loss of generality, suppose $M_2 \bullet \hat{X} \neq 0$. Then $\hat{y}_2 = 0$ by KKT condition (56). So when solving (SD), we can equivalently clamp y_2 to 0 and optimize only in y_0 and y_1. This corresponds to solving (SP) by ignoring the constraint (47). By [38], all extreme points of the new feasible region of (SP) has rank 1, and so (SP) must have an optimal solution with rank 1.

Case 2: $M_1 \bullet \hat{X} = M_2 \bullet \hat{X} = 0$ and $\operatorname{rank}(\hat{X}) \neq 2$. Let $r = \operatorname{rank}(\hat{X})$. Obviously $r > 0$ since $I_{00}\hat{X} = 1$. If $r = 1$, then (SP) already has a rank-one solution and (QP) is solved. So we only need to consider the case $r \geq 3$. By Proposition 6 with $\delta_1 = \delta_2 = 0$, there is a rank-one decomposition of $\hat{X}$ satisfying

$$\hat{X} = \mathbf{x}_1\mathbf{x}_1' + \mathbf{x}_2\mathbf{x}_2' + \ldots + \mathbf{x}_r\mathbf{x}_r' \tag{57}$$

$$M_1 \bullet \mathbf{x}_i\mathbf{x}_i' = 0, \quad for\ i = 1, \ldots, r \tag{58}$$

$$M_2 \bullet \mathbf{x}_i\mathbf{x}_i' = 0, \quad for\ i = 1, \ldots, r-2. \tag{59}$$

By Proposition 3, we have $Z(\mathbf{x}_1\mathbf{x}_1') = \mathbf{0}$. Let $\mathbf{x}_1 = (t_1, \mathbf{u}_1', \mathbf{v}_1')'$. Then

$$(58) \quad \Rightarrow \quad -s_1^2 + \|\mathbf{u}_1\|^2 = 0 \tag{60}$$

$$(59) \quad \Rightarrow \quad -s_1^2 + \|\mathbf{v}_1\|^2 = 0. \tag{61}$$

So if $s_1 = 0$ then $\mathbf{u}_1 = \mathbf{v}_1 = 0$, which means $\mathbf{x}_1 = \mathbf{0}$. Contradiction. So $s_1 \neq 0$ and we can easily see that $\hat{X}_1 := \mathbf{x}_1\mathbf{x}_1'/s_1^2$ satisfies the KKT conditions (55) and (56), together with $I_{00}\hat{X}_1 = 1$. Hence $\mathbf{x}_1\mathbf{x}_1'/s_1^2$ is a rank-one optimal solution to (SP) and $\mathbf{x}_1/s_1$ is an optimal solution to (QP).

Case 3: $M_1 \bullet \hat{X} = M_2 \bullet \hat{X} = 0$, $\operatorname{rank}(\hat{X}) = 2$, but P3 does not hold. By Proposition 4, there must be a rank-one decomposition $\hat{X} = \mathbf{x}_1\mathbf{x}_1' + \mathbf{x}_2\mathbf{x}_2'$ such that

$$M_1 \bullet (\mathbf{x}_1\mathbf{x}_1') = M_1 \bullet (\mathbf{x}_2\mathbf{x}_2') = 0. \tag{62}$$

So the failure of P3 implies

$$M_2 \bullet \mathbf{x}_1\mathbf{x}_1' = M_2 \bullet \mathbf{x}_2\mathbf{x}_2' = 0, \tag{63}$$

because $M_2 \bullet \mathbf{x}_1\mathbf{x}_1' + M_2 \bullet \mathbf{x}_2\mathbf{x}_2' = M_2 \bullet \hat{X} = 0$. Using exactly the same argument as in Case 2, we conclude that s_1, the first element of $\mathbf{x}_1$, is non-zero, and $\mathbf{x}_1\mathbf{x}_1'/s_1^2$ is a rank-one optimal solution to (SP). Obviously, $\mathbf{x}_2\mathbf{x}_2'/s_2^2$ is also a rank-one solution to (SP), where s_2 is the first element of $\mathbf{x}_2$.

Case 4: $M_1 \bullet \hat{X} = M_2 \bullet \hat{X} = 0$, $\operatorname{rank}(\hat{X}) = 2$, $M_1 \bullet (\mathbf{x}_1\mathbf{x}_1') = M_1 \bullet (\mathbf{x}_2\mathbf{x}_2') = 0$, $(M_2 \bullet \mathbf{x}_1\mathbf{x}_1')(M_2 \bullet \mathbf{x}_2\mathbf{x}_2') < 0$, and $\operatorname{rank}(\hat{Z}) \neq m + n - 1$. By Sylvester's inequality,

$$\operatorname{rank}(\hat{Z}) + \operatorname{rank}(\hat{X}) - (m + n + 1) \leq \operatorname{rank}(\hat{Z}\hat{X}). \tag{64}$$

Now $\operatorname{rank}(\hat{X}) = 2$ and $\hat{Z}\hat{X} = \mathbf{0}$, so $\operatorname{rank}(\hat{Z}) \leq m + n - 1$. Therefore in this particular case $\operatorname{rank}(\hat{Z}) \leq m + n - 2$. So by 0.4.5(d) of [39],

$$\operatorname{rank}(\hat{X} + \hat{Z}) \leq \operatorname{rank}(\hat{X}) + \operatorname{rank}(\hat{Z}) \tag{65}$$

$$\leq 2 + (m + n - 2) = m + n. \tag{66}$$

Thus there must be a $\mathbf{y} \neq \mathbf{0}$ such that $(\hat{X} + \hat{Z})\mathbf{y} = \mathbf{0}$, and

$$\mathbf{y}'\hat{X}\mathbf{y} + \mathbf{y}'\hat{Z}\mathbf{y} = \mathbf{y}'(\hat{X} + \hat{Z})\mathbf{y} = 0. \tag{67}$$

Since both $\hat{X}$ and $\hat{Z}$ are positive semi-definite, we conclude that $\mathbf{y} \in \operatorname{Ker}(\hat{X}) \cap \operatorname{Ker}(\hat{Z})$. Now define

$$X := \hat{X} + \mathbf{y}\mathbf{y}' = \mathbf{x}_1\mathbf{x}_1' + \mathbf{x}_2\mathbf{x}_2' + \mathbf{y}\mathbf{y}'. \tag{68}$$

Obviously $\operatorname{rank}(X) = 3$ and $\hat{Z}X = 0$. Since

$$M_1 \bullet (\mathbf{x}_1\mathbf{x}_1') = M_1 \bullet (\mathbf{x}_2\mathbf{x}_2') = 0 \tag{69}$$

$$(M_2 \bullet \mathbf{x}_1\mathbf{x}_1')(M_2 \bullet \mathbf{x}_2\mathbf{x}_2') < 0, \tag{70}$$

so by Proposition 5 with $\delta_1 = \delta_2 = 0$, there must be an $\mathbf{x}$ such that X is rank-one decomposable at $\mathbf{x}$ and

$$M_1 \bullet \mathbf{x}\mathbf{x}' = 0, \quad M_2 \bullet \mathbf{x}\mathbf{x}' = 0. \tag{71}$$

Since $\hat{Z}X = \mathbf{0}$, Proposition 3 implies $\hat{Z}\mathbf{x} = \mathbf{0}$ and so $\hat{Z} \bullet \mathbf{x}\mathbf{x}' = 0$. Based on the satisfaction of the KKT conditions (55) and (56), we conclude that $\mathbf{x}\mathbf{x}'/s^2$ is a rank-one optimal solution to (SP), where s is the first element of $\mathbf{x}$. s must be non-zero because of (71) and the same argument as in Case 2. ■

D.1.2 Strong Duality

Let us denote $(A, B, \mathbf{b})$ collectively as $\Gamma := (A, B, \mathbf{b})$, and define a "Frobenius" norm on Γ as $\|\Gamma\|^2 := \|A\|_F^2 + \|B\|_F^2 + \|\mathbf{b}\|^2$. Ideally we wish to show that for any Γ, the Property $\mathcal{I}$ does not hold for some solutions to (SP) and (SD), hence strong duality holds (Theorem 3). However, this is hard. So we resort to the argument of ϵ-perturbation.

Before proceeding, we first make a very simple rewriting of (QP). Let $p = \max\{t, m, n\}$. By padding zeros if necessary, we can expand A and B into p-by-p dimensional matrices, and $\mathbf{c}$ into an p dimensional vector. Let $\mathbf{x}$ and $\mathbf{y}$ be p dimensional too. Obviously, the optimal values of (QP) and (SP) in this new problem are the same as those in the original problem, respectively. Therefore, henceforth we will only consider square matrices A and B. For notational convenience, we just call all t, m, and n as n.

Of key importance is the Danskin's theorem.

Lemma 1 (Danskin) *Suppose $f : Z \times \Omega \to \mathbb{R}$ is a continuous function, where $Z \subseteq \mathbb{R}^n$ is a compact set and $\Omega \subseteq \mathbb{R}^m$ is an open set. For any $\mathbf{z}$, $\nabla_\omega f(\mathbf{z}, \omega)$ exists and is continuous. Then the marginal function*

$$\phi(\omega) := \max_{\mathbf{z} \in Z} f(\mathbf{z}, \omega) \tag{72}$$

is continuous.

Note that Danskin's theorem does not require convexity. Let the $\mathbf{z}$ in Lemma 1 correspond to $(\mathbf{x}', \mathbf{y}')'$ in (QP), ω to Γ, Z to $\{\mathbf{x} : \|\mathbf{x}\| \le 1\} \times \{\mathbf{y} : \|\mathbf{y}\| \le 1\}$, and Ω to the whole Euclidean space. Then Lemma 1 implies that $QP(\Gamma)$ is continuous in Γ. Similarly, $SP(\Gamma)$ is continuous.

The continuity at Γ means that for any $\epsilon > 0$, there exists $\delta > 0$, such that for all $\hat{\Gamma}$ in the δ neighborhood of Γ:

$$\mathcal{B}_\delta(\Gamma) := \left\{\hat{\Gamma} : \left\|\hat{\Gamma} - \Gamma\right\| < \delta\right\}, \tag{73}$$

we have

$$\left|QP(\hat{\Gamma}) - QP(\Gamma)\right| < \epsilon, \tag{74}$$

$$\left|SP(\hat{\Gamma}) - SP(\Gamma)\right| < \epsilon. \tag{75}$$

Our key result will be the following theorem.

Theorem 4 *For any Γ and $\delta > 0$, there exists $\Gamma_\delta \in \mathcal{B}_\delta(\Gamma)$ such that strong duality holds at Γ_δ:*

$$QP(\Gamma_\delta) = SP(\Gamma_\delta). \tag{76}$$

Using Theorem 4, we can easily prove strong duality.

Corollary 1 $QP(\Gamma) = SP(\Gamma)$ *for all* Γ.

Proof: It suffices to show that for any $\epsilon > 0$,

$$|QP(\Gamma) - SP(\Gamma)| < 2\epsilon. \tag{77}$$

By continuity of QP and SP, there exists a $\delta > 0$, such that (74) and (75) hold for all $\hat{\Gamma} \in \mathcal{B}_\delta(\Gamma)$. By Theorem 4, there exists $\Gamma_\delta \in \mathcal{B}_\delta(\Gamma)$ such that (76) holds. Combining it with (74) and (75) (with $\hat{\Gamma} = \Gamma_\delta$), we obtain (77). ■

Finally we prove Theorem 4.

Proof: Clearly $\mathcal{B}_{\delta/2}(A, B, \mathbf{b})$ contains invertible matrices for any A, B, and $\delta > 0$. Arbitrarily pick two such matrices and call them A_δ and B_δ. By Theorem 3, to establish (76) it suffices to show that the corresponding (SP) and (SD) problems at (A_δ, B_δ) have a pair of optimal solutions $\hat{X}$ and $(\hat{Z}, \hat{y}_0, \hat{y}_1, \hat{y}_2)$ which do not satisfy Property $\mathcal{I}$. We will focus on the second condition: $\mathrm{rank}(\hat{Z}) = 2n - 1$.

If $\mathrm{rank}(\hat{Z}) \neq 2n - 1$, then by Theorem 3 strong duality holds at $\Gamma_\delta := (A_\delta, B_\delta, \mathbf{b})$. Otherwise suppose $\mathrm{rank}(\hat{Z}) = 2n - 1$. Noting (51), we have

$$\hat{Z} = M_0 - \hat{y}_0 I_{00} + \hat{y}_1 M_1 + \hat{y}_2 M_2 = \begin{pmatrix} -\hat{y}_0 - \hat{y}_1 - \hat{y}_2 & -\mathbf{b}' \\ -\mathbf{b} & R \end{pmatrix}, \tag{78}$$

$$\text{where} \quad R = \begin{pmatrix} \hat{y}_1 I - A_\delta' A_\delta & -A_\delta' B_\delta \\ -B_\delta' A_\delta & \hat{y}_2 I - B_\delta' B_\delta \end{pmatrix}. \tag{79}$$

Note that for any given y_1 and y_2, (SD) maximizes y_0 subject to $\hat{Z} \succeq \mathbf{0}$. By Proposition 7, we know that

$$2n - 1 = \mathrm{rank}(\hat{Z}) = \mathrm{rank}(R). \tag{80}$$

Denote $P = \hat{y}_1 I - A_\delta' A_\delta$ and $Q = \hat{y}_2 I - B_\delta' B_\delta$. Then by Proposition 8, we have $\mathrm{rank}(P) + \mathrm{rank}(Q) = 2n - 1$ or $2n$. Now we discuss three cases.

Case 1: $\mathrm{rank}(P) = n$ and $\mathrm{rank}(Q) = n-1$. By Schur complement, we have $Q \succeq B_\delta' A_\delta P^{-1} A_\delta' B_\delta$. So by Exercise 4.3.14 of [39],

$$\lambda_{\min}(Q) \geq \lambda_{\min}(B_\delta' A_\delta P^{-1} A_\delta' B_\delta), \tag{81}$$

where $\lambda_{\min}$ stands for the smallest eigenvalue. Since A_δ and B_δ are both invertible, $B_\delta' A_\delta P^{-1} A_\delta' B_\delta$ must be positive definite and its smallest eigenvalue is strictly positive. But $\mathrm{rank}(Q) = n - 1$, meaning the minimum eigenvalue of Q is 0. So contraction with (81).

Case 2: $\mathrm{rank}(P) = n - 1$ and $\mathrm{rank}(Q) = n$. Same argument as for Case 1.

Case 3: $\mathrm{rank}(P) = \mathrm{rank}(Q) = n$. Since $\mathrm{rank}(R) = 2n-1$, R must have an eigen-vector $\mathbf{u}_0$ whose corresponding eigen-value is 0. In fact $\mathbf{u}_0$ is unique up to negation. By Proposition 7, $\mathbf{b} \in \mathrm{Im}(R)$, so $\mathbf{b}'\mathbf{u}_0 = 0$. Now perturb the $\mathbf{b}$ in Z in the direction of $\mathbf{u}_0$:

$$\hat{Z}(t) = \begin{pmatrix} -\hat{y}_0(t) - \hat{y}_1(t) - \hat{y}_2(t) & -\mathbf{b}' - t\mathbf{u}_0' \\ -\mathbf{b} - t\mathbf{u}_0 & R(t) \end{pmatrix}, \quad t \in \mathbb{R}, \tag{82}$$

where $\hat{y}_i(t)$ are the optimal solutions for $SD(A_\delta, B_\delta, \mathbf{b} + t\mathbf{u}_0)$ and $R(t)$ uses $\hat{y}_i(t)$. Denote $P(t) = \hat{y}_1(t) I - A_\delta' A_\delta$ and $Q(t) = \hat{y}_2(t) I - B_\delta' B_\delta$. If there exists $t \in (-\delta/2, \delta/2)$ such that $\mathrm{rank}(\hat{Z}(t)) \neq 2n-1$, then $(A_\delta, B_\delta, \mathbf{b}+t\mathbf{u}_0)$ is the Γ_δ in Theorem 4. Otherwise, $\mathrm{rank}(\hat{Z}(t)) = 2n-1$ for all $|t| < \delta/2$ and by the same argument as in Case 1 and 2, we conclude that $\mathrm{rank}(P(t)) = \mathrm{rank}(Q(t)) = n$, $\forall\ t$. Since $\mathrm{rank}(R(t)) = \mathrm{rank}(\hat{Z}(t)) = 2n - 1$, $R(t)$ must have an eigen-vector $\mathbf{u}(t)$ whose corresponding eigen-value is 0. Clearly $\mathbf{u}(t)$ is unique up to the sign, and we can set $\mathbf{u}(0) = \mathbf{u}_0$. By Proposition 7, $\mathbf{b} + t\mathbf{u}_0$ must be in the range of $R(t)$. If we can show that $\mathbf{u}(t) = (1 + ct)\mathbf{u}_0 + \mathbf{o}(t)$ where $\lim_{t\to 0} \mathbf{o}(t)/t = \mathbf{0}$ and $c \in \mathbb{R}$ is independent of t, then

$$0 = (\mathbf{b} + t\mathbf{u}_0)'\mathbf{u}(t) = (\mathbf{b} + t\mathbf{u}_0)'((1 + ct)\mathbf{u}_0 + \mathbf{o}(t)) = t + ct^2 + \mathbf{b}'\mathbf{o}(t) + t\mathbf{u}_0'\mathbf{o}(t). \tag{83}$$

Dividing both sides by t and driving t to 0, we get $0 = 1 + 0 + 0 + 0$. Contradiction.

To show $\mathbf{u}(t) = (1 + ct)\mathbf{u}_0 + \mathbf{o}(t)$, we need to analyze the gradient of $\mathbf{u}(t)$ at $t = 0$. First we show $\hat{y}_i(t)$ is differentiable in t at $t = 0$ for $i = 1, 2$. Since $\mathrm{rank}(P(t)) = \mathrm{rank}(Q(t)) = n$ and

$\text{rank}(R(t)) = 2n-1$, we have $0 = \det(R(t)) = \det(P(t)) \cdot \det(Q(t) - B_\delta' A_\delta P(t)^{-1} A_\delta' B_\delta)$. In conjunction with Schur complement, we get

$$\hat{y}_2(t) = \lambda_{\max}\left(B_\delta' B_\delta + B_\delta' A_\delta (\hat{y}_1(t) I - A_\delta' A_\delta)^{-1} A_\delta' B_\delta\right), \tag{84}$$

$$\hat{y}_1(t) = \lambda_{\max}\left(A_\delta' A_\delta + A_\delta' B_\delta (\hat{y}_2(t) I - B_\delta' B_\delta)^{-1} B_\delta' A_\delta\right). \tag{85}$$

So a larger $\hat{y}_1(t)$ implies a smaller $\hat{y}_2(t)$ and a smaller $\hat{y}_1(t)$ implies a larger $\hat{y}_2(t)$. By Proposition 7,

$$(\hat{y}_1(t), \hat{y}_2(t)) = \underset{y_1, y_2}{\text{argmin}}\, y_1 + y_2 + (\mathbf{b} + t\mathbf{u}_0)' \begin{pmatrix} y_1 I - A_\delta' A_\delta & -A_\delta' B_\delta \\ -B_\delta' A_\delta & y_2 I - B_\delta' B_\delta \end{pmatrix}^\dagger (\mathbf{b} + t\mathbf{u}_0). \tag{86}$$

In general, pseudo-inverse is not even continuous. However, since we know that $\text{rank}(R(t)) = 2n-1$ (constant rank), so the pseudo-inverse is differentiable in $R(t)$ [40]. So $\hat{y}_1(t)$ and $\hat{y}_2(t)$ are differentiable in t at $t = 0$.

By Theorem 1 of [41], we know there exists a choice of the sign for $\mathbf{u}(t)$ which satisfies

$$\left.\frac{\partial \mathbf{u}(t)}{\partial t}\right|_{t=0} = \mathbf{u}_0 \sum_{ij} A_{ij} \left.\frac{\partial R_{ij}(t)}{\partial t}\right|_{t=0}, \quad \text{where} \quad A = -R(0)^\dagger \tag{87}$$

$$= \mathbf{u}_0 \left(\hat{y}_1'(0) \sum_{i=1}^{n} A_{ii} + \hat{y}_2'(0) \sum_{i=n+1}^{2n} A_{ii}\right). \tag{88}$$

Setting $c := \hat{y}_1'(0) \sum_{i=1}^n A_{ii} + \hat{y}_2'(0) \sum_{i=n+1}^{2n} A_{ii}$ yields $\mathbf{u}(t) = (1 + ct)\mathbf{u}_0 + \mathbf{o}(t)$. ■

D.1.3 Recovering the optimal solution

With the guarantee of strong duality, an algorithm is needed to recover a rank-one optimal solution to (SP) when given an optimal dual solution $\hat{Z}$ to (SD). By the KKT condition, all we need is two vectors $\mathbf{x}$ and $\mathbf{y}$ satisfying:

$$\mathbf{z}' \hat{Z} \mathbf{z} = 0, \quad \|\mathbf{x}\| \le 1, \quad \|\mathbf{y}\| \le 1, \tag{89}$$

where $\mathbf{z} = (1, \mathbf{x}', \mathbf{y}')'$. Note this is a necessary and sufficient condition for optimal $\mathbf{x}$ and $\mathbf{y}$. Since $\hat{Z}$ is positive semi-definite, $\mathbf{z}$ must be in the null space of $\hat{Z}$. Suppose $\text{Ker}(\hat{Z})$ is spanned by $(\mathbf{g}_1, \ldots, \mathbf{g}_k)$. Let

$$G = (\mathbf{g}_1, \ldots, \mathbf{g}_k) = \begin{pmatrix} G_0 \\ G_X \\ G_Y \end{pmatrix}. \tag{90}$$

Then it suffices to find $\boldsymbol{\alpha} \in \mathbb{R}^k$ such that $|G_0 \boldsymbol{\alpha}| = 1$, $\|G_X \boldsymbol{\alpha}\| = 1$, and $\|G_Y\| \boldsymbol{\alpha} = 1$. To this end, we only need to find $\boldsymbol{\alpha}$ satisfying

$$\boldsymbol{\alpha}'(G_X' G_X - G_0' G_0)\boldsymbol{\alpha} = 0 \tag{91}$$

$$\boldsymbol{\alpha}'(G_Y' G_Y - G_0' G_0)\boldsymbol{\alpha} = 0 \tag{92}$$

$$G_0 \boldsymbol{\alpha} \neq 0, \tag{93}$$

and then scale it properly. In the sequel, we will first find $\boldsymbol{\alpha}$ which satisfies the first two conditions and then show how to satisfy the last one. Denote $\tilde{S} = G_X' G_X - G_0' G_0$ and $\tilde{T} = G_Y' G_Y - G_0' G_0$. Let their algebraically smallest eigenvalues be s_X and s_Y, and define $s = 1 - \min(s_X, s_Y)$. Then $S := \tilde{S} + sI$ and $T := \tilde{T} + sI$ must be positive definite, and $\boldsymbol{\alpha}$ only needs to satisfy

$$\boldsymbol{\alpha}' S \boldsymbol{\alpha} = s \boldsymbol{\alpha}' \boldsymbol{\alpha} \tag{94}$$

$$\boldsymbol{\alpha}' T \boldsymbol{\alpha} = s \boldsymbol{\alpha}' \boldsymbol{\alpha} \tag{95}$$

$$G_0 \boldsymbol{\alpha} \neq 0. \tag{96}$$

Denote $\hat{\boldsymbol{\alpha}} = \boldsymbol{\alpha} / \|\boldsymbol{\alpha}\|$, then it is equivalent to

$$\hat{\boldsymbol{\alpha}}' S \hat{\boldsymbol{\alpha}} = s \tag{97}$$

$$\hat{\boldsymbol{\alpha}}' T \hat{\boldsymbol{\alpha}} = s \tag{98}$$

$$G_0 \hat{\boldsymbol{\alpha}} \neq 0. \tag{99}$$

Because both S and T are positive semidefinite, by [39, Corollary 4.6.12] there exists a nonsingular matrix R such that $RSR' = I$ and RTR' is real diagonal. In fact this R can be constructed analytically. Let S have eigen-decomposition $S = UDU'$ where D is diagonal. Denote $H = U\sqrt{D}U'$ and let HTH have eigen-decomposition $HTH = V\Lambda V'$. Then R can be simply chose as $R = V'H^{-1} = VUD^{-1/2}U'$. Let RTR' be $\operatorname{diag}\{\sigma_i\}$. Denote $\boldsymbol{\beta} = R\hat{\boldsymbol{\alpha}}$, then $\boldsymbol{\beta}$ only needs to satisfy

$$\boldsymbol{\beta}'\boldsymbol{\beta} = s \tag{100}$$

$$\boldsymbol{\beta}'\Sigma\boldsymbol{\beta} = s \tag{101}$$

$$G_0 R^{-1} \boldsymbol{\beta} \neq 0. \tag{102}$$

It is easy to find a $\boldsymbol{\beta}$ which satisfies the first two constraints, because it is guaranteed that there exists a $\boldsymbol{\beta}$ which satisfies all the three conditions. Once we get such a $\boldsymbol{\beta}$, suppose $G_0 R^{-1}\boldsymbol{\beta} = 0$. Then we can flip the sign of one of its nonzero components. If its product with $G_0 R^{-1}$ is still 0, then it means the corresponding entry in $G_0 R^{-1}$ is 0. But $G_0 R^{-1}$ cannot be straight 0 because that would imply G_0 is a zero vector which violates the assumption that G is the basis of $\operatorname{Ker}(\hat{Z})$. Therefore we can always find a $\boldsymbol{\beta}$ which satisfies (100) to (102).

D.2 Preliminaries in Matrix Analysis

D.2.1 Matrix Rank-one decomposition

Let X be a n-by-n positive semi-definite matrix with $\operatorname{rank}(X) = r$. Then a set of r vectors $\{\mathbf{x}_1, \ldots, \mathbf{x}_r\}$ in $\mathbb{R}^n$ is called a rank-one decomposition of X if $X = \sum_{i=1}^r \mathbf{x}_i \mathbf{x}_i'$.

It is noteworthy that the $\mathbf{x}_i$'s are not necessarily orthogonal to each other ($\mathbf{x}_i'\mathbf{x}_j = 0$ for $i \neq j$), but they must be linearly independent. This leads to the following useful result.

Proposition 3 *Suppose $ZX = \mathbf{0}$ and $\{\mathbf{x}_1, \ldots, \mathbf{x}_r\}$ is a rank-one decomposition of X. Then $Z\mathbf{x}_i = \mathbf{0}$, $\forall\, i$.*

Proof: Denote $\mathbf{y}_i := Z\mathbf{x}_i$. Suppose otherwise $\mathbf{y}_1 \neq \mathbf{0}$. Since $ZX = \mathbf{0}$, we have

$$\mathbf{0} = X'Z'\mathbf{y}_1 = \sum_{i=1}^r \mathbf{x}_i\mathbf{x}_i' Z'\mathbf{y}_1 = \sum_{i=1}^r (\mathbf{y}_i'\mathbf{y}_1)\mathbf{x}_i. \tag{103}$$

Since $\mathbf{y}_1 \neq \mathbf{0}$, this violates the linear independence of $\mathbf{x}_1, \ldots, \mathbf{x}_r$. ■

X is called rank-one decomposable at a vector $\mathbf{x}_1$ if there exist other $r-1$ vectors $\mathbf{x}_2, \ldots, \mathbf{x}_r$ such that $X = \sum_{i=1}^r \mathbf{x}_i\mathbf{x}_i'$.

The following three theorems play an important role in our proof.

Proposition 4 (Corollary 4 of [42]) *Suppose $X \in \mathcal{S}_+^{n\times n}$ with $\operatorname{rank}(X) = r$. $Z \in \mathcal{S}^{n\times n}$ and $Z \bullet X \geq 0$. Then there must be a rank-one decomposition of $X = \mathbf{x}_1\mathbf{x}_1' + \ldots + \mathbf{x}_r\mathbf{x}_r'$ such that $Z \bullet (\mathbf{x}_i\mathbf{x}_i') = Z \bullet X/r$ for all $i = 1, \ldots, r$.*

Proposition 5 (Lemma 3.3 of [37]) *Suppose $X \in \mathcal{S}_+^{n\times n}$ with rank $r \geq 3$. $A_1, A_2 \in \mathcal{S}^{n\times n}$. Let $\{\mathbf{x}_1, \ldots, \mathbf{x}_r\}$ be a rank-one decomposition of X. If*

$$A_1 \bullet \mathbf{x}_1\mathbf{x}_1' = A_1 \bullet \mathbf{x}_2\mathbf{x}_2' = \delta_1 \tag{104}$$

$$(A_2 \bullet \mathbf{x}_1\mathbf{x}_1' - \delta_2)(A_2 \bullet \mathbf{x}_2\mathbf{x}_2' - \delta_2) < 0, \tag{105}$$

then there is a vector $\mathbf{y} \in \mathbb{R}^n$ such that X is rank-one decomposable at $\mathbf{y}$ and

$$A_1 \bullet \mathbf{y}\mathbf{y}' = \delta_1, \qquad A_2 \bullet \mathbf{y}\mathbf{y}' = \delta_2. \tag{106}$$

Proposition 6 (Theorem 3.4 of [37]) *Suppose $X \in \mathcal{S}_+^{n\times n}$ with rank $r \geq 3$. $A_1, A_2 \in \mathcal{S}^{n\times n}$. If*

$$A_1 \bullet X = \delta_1, \qquad A_2 \bullet X = \delta_2, \tag{107}$$

then X has a rank-one decomposition $\{\mathbf{x}_1, \ldots, \mathbf{x}_r\}$ such that

$$A_1 \bullet \mathbf{x}_i\mathbf{x}_i' = \delta_1/r \quad for\ i = 1, \ldots, r, \tag{108}$$

$$A_2 \bullet \mathbf{x}_i\mathbf{x}_i' = \delta_2/r \quad for\ i = 1, \ldots, r-2. \tag{109}$$

D.2.2 Bounding the rank of block matrices

Proposition 7 *Let $X \in \mathcal{S}_+^{n\times n}$ and $\mathbf{b} \in im(X)$. Define*

$$Y(c) = \begin{pmatrix} c & \mathbf{b}' \\ \mathbf{b} & X \end{pmatrix}, \qquad c \in \mathbb{R}. \tag{110}$$

Suppose $Y(c) \succeq \mathbf{0}$ and $\text{rank}(X) = r$. Then

$$\text{rank}(Y(c)) \in \{r, r+1\}. \tag{111}$$

Furthermore, if c^* is the minimum value such that $Y(c) \succeq \mathbf{0}$:

$$c^* = \operatorname*{arginf}_{c:Y(c)\succeq \mathbf{0}} c, \tag{112}$$

then we have $\text{rank}(Y(c^*)) = r$.

Finally, if $\mathbf{b} \notin im(X)$, then $Y(c) \succeq \mathbf{0}$ cannot hold for any $c \in \mathbb{R}$.

Proof: Since adding rows and columns to a matrix will not decrease its rank, so obviously $\text{rank}(Y(c)) \geq \text{rank}(X) = r$. To show $\text{rank}(Y(c)) \leq r+1$, let the eigenvalues of X and $Y(c)$ be $\lambda_1, \ldots, \lambda_n$ and $\hat{\lambda}_1, \ldots, \hat{\lambda}_{n+1}$, both in increasing order. Then by Theorem 4.3.8 of [39], we have

$$\hat{\lambda}_1 \leq \lambda_1 \leq \hat{\lambda}_2 \leq \lambda_2 \leq \ldots \hat{\lambda}_n \leq \lambda_n \leq \hat{\lambda}_{n+1}. \tag{113}$$

Since $\text{rank}(X) = r$ and $X \in \mathcal{S}_+^{n\times n}$, so $\lambda_1 = \ldots = \lambda_{n-r} = 0$. As $Y(c) \succeq \mathbf{0}$, we have

$$0 \leq \hat{\lambda}_1 \leq \ldots \leq \hat{\lambda}_{n-r} \leq 0. \tag{114}$$

Therefore $\text{rank}(Y(c)) \leq (n+1) - (n-r) = r+1$.

As for the second part, we can actually compute c^* explicitly. $Y(c) \succeq \mathbf{0}$ if and only if $(\alpha, \mathbf{u}')Y(c)(\alpha, \mathbf{u}')' \geq 0$ for all $\alpha \in \mathbb{R}$ and $\mathbf{u} \in \mathbb{R}^n$, *i.e.*

$$c\alpha^2 + 2\alpha\mathbf{b}'\mathbf{u} + \mathbf{u}'X\mathbf{u} \geq 0, \quad \forall\, \alpha \in \mathbb{R}, \mathbf{u} \in \mathbb{R}^n. \tag{115}$$

If $\alpha = 0$, this must hold true since $X \succeq \mathbf{0}$. Otherwise,

$$c^* = \max_{\alpha\neq 0,\mathbf{u}} \frac{-\mathbf{u}'X\mathbf{u} - 2\alpha\mathbf{b}'\mathbf{u}}{\alpha^2} \tag{116}$$

$$= \max_{\mathbf{z}} -\mathbf{z}'X\mathbf{z} - 2\mathbf{b}'\mathbf{z} \tag{117}$$

$$= \begin{cases} \mathbf{b}'X^\dagger\mathbf{b} & \text{if } \mathbf{b} \in im(X) \\ \infty & \text{if } \mathbf{b} \notin im(X) \end{cases}, \tag{118}$$

where $X^\dagger$ is the pseudo-inverse of X. To prove $\text{rank}(Y(c^*)) = r$, it suffices to show that $\text{Ker}(Y(c^*)) = n - r + 1$. Towards this end first note

$$Y(c^*)\begin{pmatrix} -1 \\ X^\dagger\mathbf{b} \end{pmatrix} = \begin{pmatrix} \mathbf{b}'X^\dagger\mathbf{b} & \mathbf{b}' \\ \mathbf{b} & X \end{pmatrix}\begin{pmatrix} -1 \\ X^\dagger\mathbf{b} \end{pmatrix} \tag{119}$$

$$= \begin{pmatrix} 0 \\ -\mathbf{b} + XX^\dagger\mathbf{b} \end{pmatrix} = \mathbf{0}. \tag{120}$$

where the last step also used $\mathbf{b} \in im(X)$. Hence $\begin{pmatrix} -1 \\ X^\dagger\mathbf{b} \end{pmatrix} \in \text{Ker}(Y(c^*))$.

Furthermore, $\mathrm{rank}(X) = r$ implies there are $n - r$ linearly independent vectors $\mathbf{u}_1, \ldots, \mathbf{u}_{n-r} \in \mathrm{Ker}(X)$. As $\mathbf{b} \in im(X)$, so $\mathbf{b}'\mathbf{u}_i = 0$ for all i. Therefore

$$Y(c^*) \begin{pmatrix} 0 \\ \mathbf{u}_i \end{pmatrix} = \begin{pmatrix} c^* & \mathbf{b}' \\ \mathbf{b} & X \end{pmatrix} \begin{pmatrix} 0 \\ \mathbf{u}_i \end{pmatrix} = \begin{pmatrix} \mathbf{b}'\mathbf{u}_i \\ X\mathbf{u}_i \end{pmatrix} = \mathbf{0}. \tag{121}$$

Clearly $\begin{pmatrix} -1 \\ X^\dagger \mathbf{b} \end{pmatrix}, \begin{pmatrix} 0 \\ \mathbf{u}_1 \end{pmatrix}, \ldots, \begin{pmatrix} 0 \\ \mathbf{u}_{n-r} \end{pmatrix}$ are linearly independent, so $\mathrm{Ker}(Y(c^*)) \geq n-r+1$, *i.e.* $\mathrm{rank}(Y(c^*)) \leq r$.

Finally, it is obvious from (118) that no $c \in \mathbb{R}$ makes $Y(c) \succeq \mathbf{0}$ if $\mathbf{b} \notin im(X)$. ■

Proposition 8 *Let P, Q, R be n-by-n matrices, and*

$$Z = \begin{pmatrix} P & R \\ R' & Q \end{pmatrix}. \tag{122}$$

Suppose $Z \succeq \mathbf{0}$ and $\mathrm{rank}(Z) = 2n-1$. Denote $r = \mathrm{rank}(P)$ and $s = \mathrm{rank}(Q)$. Then $r + s \in \{2n-1, 2n\}$.

Proof: Let $\mathrm{Ker}(P)$ be spanned by $\mathbf{u}_1, \ldots, \mathbf{u}_{n-r}$, and $\mathrm{Ker}(Q)$ be spanned by $\mathbf{v}_1, \ldots, \mathbf{v}_{n-s}$. Denote $\hat{\mathbf{u}}_i = \begin{pmatrix} \mathbf{u}_i \\ \mathbf{0} \end{pmatrix}$ and $\hat{\mathbf{v}}_i = \begin{pmatrix} \mathbf{0} \\ \mathbf{v}_i \end{pmatrix}$. Then

$$\hat{\mathbf{u}}_i' Z \hat{\mathbf{u}}_i = \mathbf{u}_i' P \mathbf{u}_i' = 0. \tag{123}$$

Since $Z \succeq \mathbf{0}$, so $\hat{\mathbf{u}}_i \in \mathrm{Ker}(Z)$. Similarly $\hat{\mathbf{v}}_i \in \mathrm{Ker}(Z)$. Clearly $\hat{\mathbf{u}}_1, \ldots, \hat{\mathbf{u}}_{n-r}, \hat{\mathbf{v}}_1, \ldots, \hat{\mathbf{v}}_{n-s}$ are linearly independent, therefore

$$2n - 1 = \mathrm{rank}(Z) \leq 2n - (n-r) - (n-s) = r + s. \tag{124}$$

So $r + s \in \{2n-1, 2n\}$. ■

Auxiliary References

[35] Y. Nesterov. Smooth minimization of non-smooth functions. *Mathematical Programming*, 103:127–152, 2005.

[36] F. Bach, J. Mairal, and J. Ponce. Convex sparse matrix factorizations. arXiv:0812.1869v1, 2008.

[37] W. Ai and S. Zhang. Strong duality for the CDT subproblem: A necessary and sufficient condition. *SIAM Journal on Optimization*, 19:1735–1756, 2009.

[38] G. Pataki. On the rank of extreme matrices in semidefinite programs and the multiplicity of optimal eigenvalues. *Math. Oper. Res.*, 23(2):339–358, 1998.

[39] R. A. Horn and C. R. Johnson. *Matrix Analysis*. Cambridge University Press, Cambridge, 1985.

[40] G. H. Golub and V. Pereyra. The differentiation of pseudo-inverses and nonlinear least squares problems whose variables separate. *SIAM Journal on Numerical Analysis*, 10(2):413–432, 1973.

[41] J. R. Magnus. On differentiating eigenvalues and eigenvectors. *Econometric Theory*, 1(2): 179–191, 1985.

[42] J. Sturm and S. Zhang. On cones of nonnegative quadratic functions. *Mathematics of Operations Research*, 28:246–267, 2003.