1 Appendix

(include the corrections in Lemma 2 and Theorem 1)

1.1 Proof of Lemma 2

Recall our optimization problem:

$$\min_{\boldsymbol{x}\in\mathbb{R}^n} \frac{1}{2}\boldsymbol{x}^T\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}^T\boldsymbol{x} \qquad \text{subject to } \|\boldsymbol{x}\|^2 = 1, \tag{1}$$

The proof of the global minimizers is given by Lemmas 2.4 and 2.8 in [1]. Below we provide the proof of the sufficient condition for strict local minima of problem (1). This is a consequence of the second-order sufficient condition for optimality in constrained optimization (see Chapter 3 - [2]). Notice that in our case, the Hessian of the Lagrange function is $\nabla^2_{\boldsymbol{x}\boldsymbol{x}}\mathcal{L}(\boldsymbol{x},\gamma) = \boldsymbol{A} - \gamma\boldsymbol{I}$ and the Jacobian of the constraint $\boldsymbol{x}^T\boldsymbol{x} - 1 = 0$ is $\boldsymbol{J}(\boldsymbol{x}) = \boldsymbol{x}$. Let $\boldsymbol{x}_*$ be a stationary point of problem (1). Then $\boldsymbol{x}_*$ is a strict local minimum if

$$\boldsymbol{y}^T(\boldsymbol{A} - \gamma\boldsymbol{I})\boldsymbol{y} > 0 \quad \forall\, \boldsymbol{y} \text{ s.t. } \boldsymbol{y} \perp \boldsymbol{x}_* \text{ (i.e. } \boldsymbol{y}^T\boldsymbol{x}_* = 0). \tag{2}$$

Since $\boldsymbol{P}^{\perp}_{\boldsymbol{x}_*}\boldsymbol{y} = \boldsymbol{y}$ for all $\boldsymbol{y} \perp \boldsymbol{x}$, we have

$$\begin{aligned}
\boldsymbol{y}^T(\boldsymbol{A} - \gamma\boldsymbol{I})\boldsymbol{y} &= \boldsymbol{y}^T\boldsymbol{P}^{\perp}_{\boldsymbol{x}_*}(\boldsymbol{A} - \gamma\boldsymbol{I})\boldsymbol{P}^{\perp}_{\boldsymbol{x}_*}\boldsymbol{y} \\
&= \boldsymbol{y}^T\boldsymbol{P}^{\perp}_{\boldsymbol{x}_*}\boldsymbol{A}\boldsymbol{P}^{\perp}_{\boldsymbol{x}_*}\boldsymbol{y} - \gamma\boldsymbol{y}^T\boldsymbol{P}^{\perp}_{\boldsymbol{x}_*}\boldsymbol{P}^{\perp}_{\boldsymbol{x}_*}\boldsymbol{y} \\
&= \boldsymbol{y}^T(\boldsymbol{P}^{\perp}_{\boldsymbol{x}_*}\boldsymbol{A}\boldsymbol{P}^{\perp}_{\boldsymbol{x}_*} - \gamma\boldsymbol{I})\boldsymbol{y}.
\end{aligned}$$

Thus, condition (2) is equivalent to $\boldsymbol{y}^T(\boldsymbol{P}^{\perp}_{\boldsymbol{x}_*}\boldsymbol{A}\boldsymbol{P}^{\perp}_{\boldsymbol{x}_*} - \gamma\boldsymbol{I})\boldsymbol{y} > 0$, or

$$\gamma < \frac{\boldsymbol{y}^T\boldsymbol{P}^{\perp}_{\boldsymbol{x}_*}\boldsymbol{A}\boldsymbol{P}^{\perp}_{\boldsymbol{x}_*}\boldsymbol{y}}{\|\boldsymbol{y}\|^2} \quad \forall\, \boldsymbol{y} \text{ s.t. } \boldsymbol{y} \perp \boldsymbol{x}_*. \tag{3}$$

On the other hand, by the definition of λ_{n-1}, we have

$$\lambda_{n-1} = \min_{\boldsymbol{y}\perp\boldsymbol{x}_*} \frac{\boldsymbol{y}^T\boldsymbol{P}^{\perp}_{\boldsymbol{x}_*}\boldsymbol{A}\boldsymbol{P}^{\perp}_{\boldsymbol{x}_*}\boldsymbol{y}}{\|\boldsymbol{y}\|^2} = \min_{\boldsymbol{y}\perp\boldsymbol{x}_*} \frac{\boldsymbol{y}^T\boldsymbol{A}\boldsymbol{y}}{\|\boldsymbol{y}\|^2} = \min_{\substack{\boldsymbol{y}\perp\boldsymbol{x}_* \\ \|\boldsymbol{y}\|=1}} \boldsymbol{y}^T\boldsymbol{A}\boldsymbol{y}. \tag{4}$$

Combining (2), (3) and (4), we conclude $\gamma < \lambda_{n-1}$ implies $\boldsymbol{x}_*$ is a strict local minimum of problem (1).

It is noteworthy that the necessary condition for local minima of problem (1), following a similar argument, is given by $\gamma \leq \lambda_{n-1}$. However, it is possible that a strict local minimum associates with $\gamma = \lambda_{n-1}$. For example, consider the 2D-case

$$\boldsymbol{A} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}, \qquad \boldsymbol{x}_* = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}, \qquad \boldsymbol{b} = \begin{bmatrix} \sqrt{2} \\ \sqrt{2} \end{bmatrix}, \qquad \gamma = \lambda_{n-1} = 1.$$

It can be seen that the curvature of the objective function almost coincides with that of the unit sphere at $\boldsymbol{x}_*$ in the above example. The following lemma states the necessary condition for strict local minima of problem (1):

Lemma 1. *If $\boldsymbol{x}_*$ is a strict local minimum of problem (1), then either of the following condition holds*

- $\gamma < \lambda_{n-1}$
- $\boldsymbol{x}_*^T\boldsymbol{A}\boldsymbol{x}_* > \boldsymbol{u}^T\boldsymbol{A}\boldsymbol{u} = \gamma = \lambda_{n-1}$ *and* $\boldsymbol{x}_*^T\boldsymbol{A}\boldsymbol{u} = 0$ *for* $\boldsymbol{u} = \operatorname*{argmin}_{\substack{\boldsymbol{y}\perp\boldsymbol{x}_* \\ \|\boldsymbol{y}\|=1}} \boldsymbol{y}^T\boldsymbol{A}\boldsymbol{y}$.

Proof. By definition of strict local minima, for any $\boldsymbol{x} \in \mathcal{S}^{n-1}$ such that $0 < \|\boldsymbol{x} - \boldsymbol{x}_*\| < \epsilon$ with sufficiently small $\epsilon > 0$, we have

$$\begin{aligned}
0 &< f(\boldsymbol{x}) - f(\boldsymbol{x}_*) \\
&= \Big(\frac{1}{2}\boldsymbol{x}^T\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}^T\boldsymbol{x}\Big) - \Big(\frac{1}{2}\boldsymbol{x}_*^T\boldsymbol{A}\boldsymbol{x}_* - \boldsymbol{b}^T\boldsymbol{x}_*\Big) \\
&= \Big(\frac{1}{2}\boldsymbol{x}^T\boldsymbol{A}\boldsymbol{x} - (\boldsymbol{A}\boldsymbol{x}_* - \gamma\boldsymbol{x}_*)^T\boldsymbol{x}\Big) - \Big(\frac{1}{2}\boldsymbol{x}_*^T\boldsymbol{A}\boldsymbol{x}_* - (\boldsymbol{A}\boldsymbol{x}_* - \gamma\boldsymbol{x}_*)^T\boldsymbol{x}_*\Big) \qquad (\text{since } \boldsymbol{A}\boldsymbol{x}_* - \boldsymbol{b} = \gamma\boldsymbol{x}_*) \\
&= \frac{1}{2}\Big(\boldsymbol{x}^T\boldsymbol{A}\boldsymbol{x} - 2\boldsymbol{x}^T\boldsymbol{A}\boldsymbol{x}_* + \boldsymbol{x}_*^T\boldsymbol{A}\boldsymbol{x}_* - \gamma(2\boldsymbol{x}_*^T\boldsymbol{x}_* - 2\boldsymbol{x}^T\boldsymbol{x}_*)\Big) \\
&= \frac{1}{2}\Big((\boldsymbol{x} - \boldsymbol{x}_*)^T\boldsymbol{A}(\boldsymbol{x} - \boldsymbol{x}_*) - \gamma\left\|\boldsymbol{x} - \boldsymbol{x}_*\right\|^2\Big). \qquad (\text{since } \|\boldsymbol{x}\| = \|\boldsymbol{x}_*\| = 1)
\end{aligned}$$

Denote $\boldsymbol{\delta} = \boldsymbol{x} - \boldsymbol{x}_* = \boldsymbol{\delta}_x + \boldsymbol{\delta}_\perp$, where $\boldsymbol{\delta}_x$ is collinear to $\boldsymbol{x}_*$ and $\boldsymbol{\delta}_\perp$ is orthogonal to $\boldsymbol{x}_*$. The last inequality becomes

$$\gamma < \frac{\boldsymbol{\delta}^T\boldsymbol{A}\boldsymbol{\delta}}{\|\boldsymbol{\delta}\|^2} = \frac{\boldsymbol{\delta}_x^T\boldsymbol{A}\boldsymbol{\delta}_x + 2\boldsymbol{\delta}_x^T\boldsymbol{A}\boldsymbol{\delta}_\perp + \boldsymbol{\delta}_\perp\boldsymbol{A}\boldsymbol{\delta}_\perp}{\|\boldsymbol{\delta}\|^2}. \tag{5}$$

Using the fact that $\|\boldsymbol{\delta}\|^2 = \|\boldsymbol{\delta}_x\|^2 + \|\boldsymbol{\delta}_\perp\|^2$ and

$$1 = \|\boldsymbol{x}\|^2 = \|\boldsymbol{x}_* + \boldsymbol{\delta}\|^2 = \|\boldsymbol{x}_*\|^2 + \|\boldsymbol{\delta}\|^2 + 2\boldsymbol{x}_*^T\boldsymbol{\delta} = 1 + \|\boldsymbol{\delta}_x\|^2 + \|\boldsymbol{\delta}_\perp\|^2 + 2\boldsymbol{x}_*^T\boldsymbol{\delta}_x$$

we obtain $\boldsymbol{\delta}_x = -\|\boldsymbol{\delta}_x\|\,\boldsymbol{x}_*$ and $\|\boldsymbol{\delta}_\perp\| = \sqrt{2\|\boldsymbol{\delta}_x\| - \|\boldsymbol{\delta}_x\|^2}$. Substituting back into (5) yields

$$\begin{aligned}
\gamma &< \frac{\|\boldsymbol{\delta}_x\|^2\boldsymbol{x}_*^T\boldsymbol{A}\boldsymbol{x}_* - 2\|\boldsymbol{\delta}_x\|\sqrt{2\|\boldsymbol{\delta}_x\| - \|\boldsymbol{\delta}_x\|^2}\,\boldsymbol{x}_*^T\boldsymbol{A}\boldsymbol{u} + (2\|\boldsymbol{\delta}_x\| - \|\boldsymbol{\delta}_x\|^2)\boldsymbol{u}^T\boldsymbol{A}\boldsymbol{u}}{2\|\boldsymbol{\delta}_x\|} \\
&= \frac{1}{2}\Big(\|\boldsymbol{\delta}_x\|\,\boldsymbol{x}_*^T\boldsymbol{A}\boldsymbol{x}_* - 2\sqrt{2\|\boldsymbol{\delta}_x\| - \|\boldsymbol{\delta}_x\|^2}\,\boldsymbol{x}_*^T\boldsymbol{A}\boldsymbol{u} - \|\boldsymbol{\delta}_x\|\,\boldsymbol{u}^T\boldsymbol{A}\boldsymbol{u}\Big) + \boldsymbol{u}^T\boldsymbol{A}\boldsymbol{u},
\end{aligned} \tag{6}$$

where $\boldsymbol{u}$ is the unit-length vector that is collinear to $\boldsymbol{\delta}_\perp$. Now since $\|\boldsymbol{\delta}_x\|$ can be chosen arbitrarily small and $\boldsymbol{u}$ can be chosen in any direction that is orthogonal to $\boldsymbol{x}_*$, taking $\|\boldsymbol{\delta}_x\| \to 0$ in (6) yields $\boldsymbol{u}^T\boldsymbol{A}\boldsymbol{u} \geq \gamma$ for any unit-length vector $\boldsymbol{u} \perp \boldsymbol{x}_*$. Thus, from (4), we conclude that $\lambda_{n-1} \geq \gamma$. Furthermore, if $\lambda_{n-1} = \boldsymbol{u}^T\boldsymbol{A}\boldsymbol{u} = \gamma$, then it holds that

$$\|\boldsymbol{\delta}_x\|\,\boldsymbol{x}_*^T\boldsymbol{A}\boldsymbol{x}_* - 2\sqrt{2\|\boldsymbol{\delta}_x\| - \|\boldsymbol{\delta}_x\|^2}\,\boldsymbol{x}_*^T\boldsymbol{A}\boldsymbol{u} - \|\boldsymbol{\delta}_x\|\,\boldsymbol{u}^T\boldsymbol{A}\boldsymbol{u} > 0 \quad \text{for all } \|\boldsymbol{\delta}_x\|. \tag{7}$$

Notice that if $\boldsymbol{x}_*^T\boldsymbol{A}\boldsymbol{u} > 0$, we can always choose sufficiently small $\|\boldsymbol{\delta}_x\|$ so that the second term ($\mathcal{O}(\|\boldsymbol{\delta}_x\|^{1/2})$) on the LHS of (7) dominates the other terms ($\mathcal{O}(\|\boldsymbol{\delta}_x\|)$), which in turn forces the LHS to be negative. Otherwise, if $\boldsymbol{x}_*^T\boldsymbol{A}\boldsymbol{u} < 0$, we can replace $\boldsymbol{u}$ by $-\boldsymbol{u}$ and follows the same argument to expose the contradiction. Therefore, it must hold that $\boldsymbol{x}_*^T\boldsymbol{A}\boldsymbol{u} = 0$ in the case $\boldsymbol{u}^T\boldsymbol{A}\boldsymbol{u} = \gamma$. In addition, substituting these quantities back into (7) yields $\boldsymbol{x}_*^T\boldsymbol{A}\boldsymbol{x}_* > \boldsymbol{u}^T\boldsymbol{A}\boldsymbol{u}$. □

1.2 Proof of Lemma 4

This lemma stems from the fact that the first-order derivative of the function $f(\boldsymbol{x}) = \frac{\boldsymbol{x}}{\|\boldsymbol{x}\|}$ is given by

$$\nabla f(\boldsymbol{x}) = \frac{1}{\|\boldsymbol{x}\|}\boldsymbol{I} - \frac{1}{\|\boldsymbol{x}\|^3}\boldsymbol{x}\boldsymbol{x}^T.$$

1.3 Proof of Lemma 5

We have

$$\alpha_* = \underset{\substack{\alpha > 0 \\ \alpha(\lambda_1 + \gamma) < 2}}{\operatorname{argmin}} \max_{1 \leq i \leq n-1} \frac{|1 - \alpha\lambda_i|}{1 - \alpha\gamma} \tag{8}$$

For $\gamma < \lambda$, the function $\frac{1-\alpha\lambda}{1-\alpha\gamma}$ is monotonically decreasing. Denote $f(\alpha) = \max_{1\le i\le n-1} \frac{|1-\alpha\lambda_i|}{1-\alpha\gamma}$. Consider the following three cases:
- If $1-\alpha\lambda_{n-1} \ge 1-\alpha\lambda_1 \ge 0$, then (8) becomes

$$\begin{aligned}\min_\alpha f(\alpha) &= \min_{\alpha\lambda_1\le 1} \frac{1-\alpha\lambda_{n-1}}{1-\alpha\gamma} \\ &= \begin{cases} f(\frac{1}{\lambda_1}) = \frac{\lambda_1-\lambda_{n-1}}{\lambda_1-\gamma} & \text{if } \lambda_1 > 0 \\ f(\infty) = \frac{\lambda_{n-1}}{\gamma} & \text{otherwise} \end{cases}\end{aligned}$$

- If $1-\alpha\lambda_1 \le 1-\alpha\lambda_{n-1} \le 0$, then (8) becomes

$$\min_\alpha f(\alpha) = \min_{\alpha\lambda_{n-1}\ge 1} \frac{\alpha\lambda_1 - 1}{1-\alpha\gamma} = f\Big(\frac{1}{\lambda_{n-1}}\Big) = \frac{\lambda_1-\lambda_{n-1}}{\lambda_{n-1}-\gamma}$$

- If $\begin{cases} 1-\alpha\lambda_1 \le 0 \\ 1-\alpha\lambda_{n-1} \ge 0 \end{cases}$, then (8) becomes

$$\begin{aligned}\min_\alpha f(\alpha) &= \min_{\alpha(\lambda_1+\lambda_{n-1})\le 2} \Big\{\frac{\alpha\lambda_1-1}{1-\alpha\gamma}, \frac{1-\alpha\lambda_{n-1}}{1-\alpha\gamma}\Big\} \\ &= \begin{cases} f(\frac{2}{\lambda_1+\lambda_{n-1}}) = \frac{\lambda_1-\lambda_{n-1}}{\lambda_1+\lambda_{n-1}-2\gamma} & \text{if } \alpha(\lambda_1+\lambda_{n-1}) < 2 \\ f(\infty) = \frac{\lambda_{n-1}}{\gamma} & \text{otherwise} \end{cases}\end{aligned}$$

In summary, we have
- If $\lambda_1 + \lambda_{n-1} \le 0$, then

$$\min_\alpha f(\alpha) = \min\Big\{f(\frac{1}{\lambda_1}), f(\infty)\Big\} = f(\infty)$$

- If $\lambda_1 + \lambda_{n-1} > 0$, then

$$\begin{aligned}\min_\alpha f(\alpha) &= \min\Big\{f(\frac{1}{\lambda_1}), f(\frac{1}{\lambda_{n-1}}), f(\frac{2}{\lambda_1+\lambda_{n-1}})\Big\} \\ &= f(\frac{2}{\lambda_1+\lambda_{n-1}}).\end{aligned}$$

References

[1] Danny C Sorensen, "Newton's method with a model trust region modification," *SIAM Journal on Numerical Analysis*, vol. 19, no. 2, pp. 409–426, 1982.

[2] D. P. Bertsekas, *Nonlinear programming*, Athena Scientific optimization and computation series. Athena Scientific, 1999.