Appendix

Proof of Lemma 2

Proof. This lemma is a straightforward consequence of the second-order necessary and sufficient conditions for optimality [1]. Notice that in our case, the Hessian of the Lagrangge function $\nabla^2_{\boldsymbol{xx}}\mathcal{L}(\boldsymbol{x},\gamma) = \boldsymbol{A} - \gamma\boldsymbol{I}$ and the constraint $\boldsymbol{x}^T\boldsymbol{x} - 1 = 0$ has the Jacobian $\boldsymbol{J}(\boldsymbol{x}) = \boldsymbol{x}$.

- **Necessary conditions.** If $\boldsymbol{x}_*$ is a local minimum of problem (2), then it satisfies the stationary conditions in Lemma 1 and

$$\boldsymbol{y}^T(\boldsymbol{A} - \gamma\boldsymbol{I})\boldsymbol{y} \geq 0 \quad \forall\ \boldsymbol{y} \text{ s.t. } \boldsymbol{y}^T\boldsymbol{x}_* = 0. \tag{1}$$

  Since $\boldsymbol{P}^{\perp}_{\boldsymbol{x}_*}\boldsymbol{y} = \boldsymbol{y}$ for all $\boldsymbol{y}$ orthogonal to $\boldsymbol{x}$, 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 (1) also implies

$$\boldsymbol{y}^T(\boldsymbol{P}^{\perp}_{\boldsymbol{x}_*}\boldsymbol{A}\boldsymbol{P}^{\perp}_{\boldsymbol{x}_*} - \gamma\boldsymbol{I})\boldsymbol{y} \geq 0 \quad \forall\ \boldsymbol{y} \text{ s.t. } \boldsymbol{y}^T\boldsymbol{x}_* = 0.$$

- **Sufficient conditions.** Let $\boldsymbol{x}_*$ be a stationary point of problem (2). By the same argument, we can also prove that $\boldsymbol{x}_*$ is a strict local minimum if

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

Now it is sufficient to show that (2) is equivalent to $\gamma < \lambda_{n-1}$. In fact, this can be shown by the definition of λ_{n-1}:

$$\lambda_{n-1} = \min_{\substack{\boldsymbol{u}\in\mathcal{S}^{n-1}\\ \boldsymbol{u}\neq\boldsymbol{x}_*}} \boldsymbol{u}^T\boldsymbol{P}^{\perp}_{\boldsymbol{x}_*}\boldsymbol{A}\boldsymbol{P}^{\perp}_{\boldsymbol{x}_*}\boldsymbol{u}.$$

For the global minimizer result, we refer to Lemmas 2.4 and 2.8 in [2]. $\square$

Proof of Lemma 4

Since $\frac{\partial}{\partial\boldsymbol{x}}\|\boldsymbol{x}\| = \frac{\boldsymbol{x}}{\|\boldsymbol{x}\|}$ and $\frac{\partial^2}{\partial\boldsymbol{x}\partial\boldsymbol{x}^T}\|\boldsymbol{x}\| = \frac{1}{\|\boldsymbol{x}\|}\boldsymbol{I} - \frac{1}{\|\boldsymbol{x}\|^2}\frac{\boldsymbol{x}}{\|\boldsymbol{x}\|}\boldsymbol{x}^T$, the Taylor series expansion of function $f(\boldsymbol{x}) = \frac{\boldsymbol{x}}{\|\boldsymbol{x}\|}$ is given by

$$\begin{aligned}f(\boldsymbol{x}+\boldsymbol{\delta}) &= f(\boldsymbol{x}) + \langle\nabla f(\boldsymbol{x}), \boldsymbol{\delta}\rangle + O(\|\boldsymbol{\delta}\|^2) \\ &= f(\boldsymbol{x}) + \left(\frac{1}{\|\boldsymbol{x}\|}\boldsymbol{I} - \frac{1}{\|\boldsymbol{x}\|^2}\frac{\boldsymbol{x}}{\|\boldsymbol{x}\|}\boldsymbol{x}^T\right)\boldsymbol{\delta} + O(\|\boldsymbol{\delta}\|^2) \\ &= f(\boldsymbol{x}) + \frac{1}{\|\boldsymbol{x}\|}\left(\boldsymbol{I} - \frac{\boldsymbol{x}\boldsymbol{x}^T}{\|\boldsymbol{x}\|^2}\right)\boldsymbol{\delta} + O(\|\boldsymbol{\delta}\|^2).\end{aligned}$$

Proof of Lemma 5

We have

$$\alpha_* = \operatorname*{argmin}_{\substack{\alpha>0\\ \alpha(\lambda_1+\gamma)<2}} \max_{1\leq i\leq n-1} \frac{|1-\alpha\lambda_i|}{1-\alpha\gamma} \tag{3}$$

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 (3) 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 (3) 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 (3) 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] D. P. Bertsekas, *Nonlinear programming*, Athena Scientific optimization and computation series. Athena Scientific, 1999.

[2] 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.