Solution of Problem 09-002, by ANTÔNIO FRANCISCO NETO[1] (Universidade Federal de Ouro Preto, Ouro Preto, Brazil).

We will show that the function $\mathcal{D}(\mathbf{A}) = \mathrm{Tr}(\mathbf{A}^{-2})$ is convex on the space of all positive-definite $m \times m$ symmetric matrices. This makes sense since the positive-definite matrices form an open convex subset of the linear space of all $m \times m$ symmetric matrices. The function $\mathcal{C}$ from the original problem statement is a composition of $\mathcal{D}$ with a linear function. Since a linear function does not effect convexity, this solves the original problem. (In fact, it is not too hard to see that this statement is equivalent to the original problem.) The argument below actually shows that $\mathcal{D}$ is strictly convex, whereas $\mathcal{C}$ need not be strictly convex.

Since convexity is implied by positivity of the second derivative, it suffices to show that for any symmetric positive-definite matrix $\mathbf{A}_0$ and any nonzero symmetric matrix $\mathbf{C}$, the function

$$f(t) = \mathcal{D}(\mathbf{A}_0 + \mathbf{C}t) = \mathrm{Tr}((\mathbf{A}_0 + \mathbf{C}t)^{-2})$$

satisfies $f''(0) > 0$. Let $\mathbf{A}(t) = \mathbf{A}_0 + \mathbf{C}t$. It is well known that

$$\frac{d\mathbf{A}^{-1}(t)}{dt} = -\mathbf{A}^{-1}(t)\frac{d\mathbf{A}(t)}{dt}\mathbf{A}^{-1}(t).$$

Since $\frac{d\mathbf{A}(t)}{dt} = \mathbf{C}$ is independent of t, we compute

$$\begin{aligned}\frac{d\mathbf{A}^{-2}(t)}{dt} &= -\mathbf{A}^{-2}(t)\frac{d\mathbf{A}(t)}{dt}\mathbf{A}^{-1}(t) - \mathbf{A}^{-1}(t)\frac{d\mathbf{A}(t)}{dt}\mathbf{A}^{-2}(t)\\ &= -\mathbf{A}^{-2}(t)\mathbf{C}\mathbf{A}^{-1}(t) - \mathbf{A}^{-1}(t)\mathbf{C}\mathbf{A}^{-2}(t)\end{aligned}$$

and

$$\begin{aligned}\frac{d^2}{dt^2}\mathbf{A}^{-2}(t) = 2\mathbf{A}^{-2}(t)\mathbf{C}\mathbf{A}^{-1}(t)\mathbf{C}\mathbf{A}^{-1}(t) + 2\mathbf{A}^{-1}(t)\mathbf{C}\mathbf{A}^{-2}(t)\mathbf{C}\mathbf{A}^{-1}(t)\\ + 2\mathbf{A}^{-1}(t)\mathbf{C}\mathbf{A}^{-1}(t)\mathbf{C}\mathbf{A}^{-2}(t).\end{aligned}$$

Therefore, we get

$$f''(0) = 2\ \mathrm{Tr}(\mathbf{A}_0^{-2}\mathbf{C}\mathbf{A}_0^{-1}\mathbf{C}\mathbf{A}_0^{-1}) + 2\ \mathrm{Tr}(\mathbf{A}_0^{-1}\mathbf{C}\mathbf{A}_0^{-2}\mathbf{C}\mathbf{A}_0^{-1}) + 2\ \mathrm{Tr}(\mathbf{A}_0^{-1}\mathbf{C}\mathbf{A}_0^{-1}\mathbf{C}\mathbf{A}_0^{-2}).$$

The middle term on the right-hand side is clearly positive since it is $\mathrm{Tr}(\mathbf{N}\mathbf{N}^*)$ for the nonzero matrix $\mathbf{N} = \mathbf{A}_0^{-1}\mathbf{C}\mathbf{A}_0^{-1}$. For the other two terms (which are easily seen to be equal), we

[1]E-mail: antoniofrancisco1975@gmail.com

use the Cholesky decomposition to write $\mathbf{A}_0^{-1} = \mathbf{L}\mathbf{L}^*$, where $\mathbf{L}$ is an invertible $m \times m$ lower triangular matrix. Then by cyclic invariance of the trace, we have

$$\mathrm{Tr}(\mathbf{A}_0^{-2}\mathbf{C}\mathbf{A}_0^{-1}\mathbf{C}\mathbf{A}_0^{-1}) = \mathrm{Tr}(\mathbf{A}_0^{-1}\mathbf{C}\mathbf{A}_0^{-1}\mathbf{C}\mathbf{A}_0^{-2}) = \mathrm{Tr}(\mathbf{L}^*\mathbf{A}_0^{-1}\mathbf{C}\mathbf{L}\mathbf{L}^*\mathbf{C}\mathbf{A}_0^{-1}\mathbf{L}),$$

which is $\mathrm{Tr}(\mathbf{M}\mathbf{M}^*)$ for the nonzero matrix $\mathbf{M} = \mathbf{L}^*\mathbf{A}_0^{-1}\mathbf{C}\mathbf{L}$. Hence these terms are also positive and we conclude that $f''(0) > 0$.

Essentially the same argument would show that for all positive integers k, the function $\mathcal{D}_k(\mathbf{A}) = \mathrm{Tr}(\mathbf{A}^{-k})$ is strictly convex on the space of all positive-definite $m \times m$ symmetric matrices.