Chapter II
Inequality Constraints

March 11, 2023

1 Optimality Conditions

2.1.1

Prove the normal cone is a closed convex cone.

Proof: Let $C \subseteq \mathbb{E}$ be a convex set and $\bar{x} \in C$. Then

$$N_C(\bar{x}) = \{d \in \mathbb{E} : \langle d, x - \bar{x} \rangle \leq 0\}.$$

If $d \in N_C(\bar{x})$, then obviously $\alpha d \in N_C(\bar{x})$ for $\alpha \geq 0$. Also, if $d_i \in N_C(\bar{x})$ and $d_i \to d \in \mathbb{E}$, then for a fixed $x \in C$,

$$\langle d_i, x - \bar{x} \rangle \to \langle d, x - \bar{x} \rangle,$$

and since $\langle d_i, x - \bar{x} \rangle \leq 0$, $\langle d, x - \bar{x} \rangle \leq 0$. Thus $d \in N_C(\bar{x})$ and so $N_C(\bar{x})$ is closed. Convexity is also clear.

2.1.3 Self-dual cones

Prove that each of the following cones K satisfy the relationship $N_K(0) = -K$.

- $\mathbb{R}^n_+$

  Proof: Recall that

  $$N_K(\bar{x}) = \{d \in \mathbb{E} : \langle d, x - \bar{x} \rangle \leq 0\}.$$

  Thus, $d \in N_{\mathbb{R}^n_+}(0)$ if and only if

  $$\langle d, x \rangle \leq 0 \quad \forall x \in \mathbb{R}^n_+.$$

  So, for $x = e_i$, we realize that $x_i \leq 0$ and hence $x \in -\mathbb{R}^n_+$. Conversely, for all $x \in -\mathbb{R}^n_+$ and all $d \in \mathbb{R}^n_+$ we clearly have $\langle d, x \rangle \leq 0$. Thus $N_{\mathbb{R}^n_+}(0) = -\mathbb{R}^n_+$.

- $\mathbb{S}^n_+$

  Proof: Recall that

  $$N_{\mathbb{S}^n_+}(0) = \{X \in \mathbb{S}^n : \langle X, A \rangle \leq 0\} \quad \forall A \in \mathbb{S}^n_+.$$

  So clearly if $-X \in \mathbb{S}^n_+$ then $X \in N_{\mathbb{S}^n_+}(0)$. Conversely, suppose $X \in N_{\mathbb{S}^n_+}(0)$, then

  $$a^T X a \leq 0 \quad \forall a \in \mathbb{R}^n \text{ or equivalently } \; a^T(-X)a \geq 0 \quad \forall a \in \mathbb{R}^n.$$

  Thus by definition $-X \in \mathbb{S}^n_+$.

- $K = \{x \in \mathbb{R}^n : x_1 \geq 0, x_1^2 \geq x_2^2 + \cdots + x_n^2\}$

  Proof: Let $y \in N_K(0)$. Then

  $$\begin{gathered} \inf \langle -y, x \rangle \\ x_1^2 \geq x_2^2 + \cdots + x_n^2 \end{gathered}$$

  has nonnegative optimum value. Note that if $\bar{x}$ is a local minimum of the function $f(x) := \langle -y, x \rangle$, then if $\bar{x} \neq 0$ then $\bar{x}_1 \neq 0$ as otherwise since $\bar{x}_1^2 \geq \bar{x}_2^2 + \cdots + \bar{x}_n^2$, $\bar{x} = 0$. So $y \in N_K(0)$ if and only if the following problem has optimum value at least 0,

  $$\begin{gathered} \inf -y_1 - \sum_{i=2}^{n} y_i x_i \\ 1 \geq x_2^2 + \cdots + x_n^2. \end{gathered}$$

However,

$$\sum_{i=2}^{n} y_i x_i \leq (x_2^2 + \cdots + x_n^2)^{\frac{1}{2}} (y_2^2 + \cdots + y_n^2)^{\frac{1}{2}} \leq (y_2^2 + \cdots + y_n^2)^{\frac{1}{2}}.$$

Now let $\bar{x} = (1, 0, 0, \cdots, 0)$ and so $-y_1 \geq 0$. Now let $x_i = \frac{y_i}{\sqrt{\sum_{i=2}^n y_i^2}}$ for $2 \leq i \leq n$, the above inequality holds with equality. Thus,

$$\sup \sum_{i=2}^{n} x_i y_i$$
$$1 \geq x_2^2 + \cdots + x_n^2,$$

has optimum value $(y_2^2 + \cdots + y_n^2)^{\frac{1}{2}}$. Thus, $y \in N_K(0)$ if and only if $(y_2^2 + \cdots + y_n^2)^{\frac{1}{2}} \leq -y_1$ or equivalently $-y \in N_K(0)$.

2.1.7

Suppose a convex function $g : [0, 1] \to \mathbb{R}$ satisfies $g(0) = 0$. Prove the function $t \in (0, 1] \mapsto g(t)/t$ is nondecreasing. Hence prove that for a convex function $f : C \to \mathbb{R}$ and points $\bar{x}, x \in C \subseteq \mathbb{E}$, the quotient $(f(\bar{x} + t(x - \bar{x}) - f(\bar{x}))/t$ is nondecreasing as a function of t in $(0, 1]$, and complete the proof of Proposition 2.1.2.

Proof: Note that g is convex and thus $g(ts) \leq t g(s)$ for all $t, s \in [0, 1]$. Thus if $t, s \neq 0$ then

$$\frac{g(ts)}{ts} \leq \frac{g(s)}{s},$$

and this means that g is nondecreasing. The rest is clear as $t \in (0, 1] \mapsto (f(\bar{x} + t(x - \bar{x})) - f(\bar{x})/t$ is convex. Hence, since $f'(\bar{x}, x - \bar{x}) = \lim_{t \to 0} (f(\bar{x} + t(x - \bar{x})) - f(\bar{x})/t \geq 0$, we conclude that $(f(\bar{x} + 1(x - \bar{x})) - f(\bar{x})/1 \geq 0$ or equivalently $f(x) \geq f(\bar{x})$.

2.1.10

- Prove the function $f : \mathbb{S}^n_{++} \to \mathbb{R}$ defined by $f(X) = \mathrm{Tr}(X^{-1})$ is differentiable on $\mathbb{S}^n_{++}$.
- Define a function $f : \mathbb{S}^n_{++} \to \mathbb{R}$ by $f(X) = \log \det(X)$. Prove $\nabla f(I) = I$. Deduce $\nabla f(X) = X^{-1}$ for any X in $\mathbb{S}^n_{++}$.

Proof:

- Let $f(X) = \mathrm{Tr}(X^{-1})$. Note that for $H \in \mathbb{S}^n$ and small enough $|t|$,

  $$\mathrm{Tr}((X + tH)^{-1}) = \mathrm{Tr}(X^{-\frac{1}{2}}(I + tX^{-\frac{1}{2}} H X^{-\frac{1}{2}})^{-1} X^{-\frac{1}{2}}) =$$
  $$\mathrm{Tr}(X^{-\frac{1}{2}}(I - tX^{-\frac{1}{2}} H X^{-\frac{1}{2}} + O(t^2)) X^{-\frac{1}{2}}) = \mathrm{Tr}(X^{-1}) - t\,\mathrm{Tr}(X^{-2}H) + O(t^2).$$

  Hence, $\nabla f(X)[H] = \mathrm{Tr}(-X^{-2}H)$. Thus, $\nabla f(X) = -X^{-2}$.

- Note that for $H \in \mathbb{S}^n$,

  $$\log \det(X + tH) = \log(1 + t\,\mathrm{Tr}(X^{-\frac{1}{2}} H X^{-\frac{1}{2}}) + O(t^2)) + \log \det(X)$$

  So,

  $$\nabla f(X)(H) = \lim_{t \to 0} \frac{1}{t} \log(1 + t\,\mathrm{Tr}(X^{-\frac{1}{2}} H X^{-\frac{1}{2}}) + O(t^2))$$
  $$= \lim_{t \to 0} \frac{\mathrm{Tr}(X^{-\frac{1}{2}} H X^{-\frac{1}{2}}) + O(t)}{1 + t\,\mathrm{Tr}(X^{-\frac{1}{2}} H X^{-\frac{1}{2}}) + O(t^2)} = \mathrm{Tr}(X^{-\frac{1}{2}} H X^{-\frac{1}{2}}) = \mathrm{Tr}(X^{-1}H).$$

  Thus $\nabla f(X) = X^{-1}$.

  Side: This is also an immediate consequence of chain rule via using the fact $\nabla \det(X) = adj(X)$.

2.1.10 Matrix completion

For a set $\Delta \subseteq \{(i,j) : 1 \le i \le j \le n\}$, suppose the subspace $L \subseteq \mathbb{S}^n$ of matrices with (i,j)th entry of zero for all (i,j) in Δ satisfies $L \cap \mathbb{S}^n_{++} \neq \emptyset$. By considering the problem (for $C \in \mathbb{S}^n_{++}$)

$$\inf\{\langle C, X\rangle - \log\det X : X \in L \cap \mathbb{S}^n_{++}\}, \tag{1}$$

prove there exists a matrix X in $L \cap \mathbb{S}^n_{++}$ with $C - X^{-1}$ having (i,j)th entry of zero for all (i,j) not in Δ.

We now the function $X \in \mathbb{S}^n_{++} \mapsto \langle C, X\rangle - \log\det(X)$ has compact level sets. Now let $A_{i,j} \in \mathbb{S}^n$ be symmetric matrices with (i,j)th and (j,i)th entry equals to 1 and 0 elsewhere. Then

$$L = \{X \in \mathbb{S}^n : \langle A_{i,j}, X\rangle = 0 \; \forall \; (i,j) \in \Delta\},$$

which is a closed subspace of $\mathbb{S}^n$. So the level sets of 1 are also compact and thus there exists a global minimizer $\bar{X}$ in $L \cap \mathbb{S}^n_{++}$. So due to *"First order conditions for linear constraints"* there exists $y_{i,j}$ for all $(i,j) \in \Delta$ such that

$$C - \bar{X}^{-1} = \sum_{(i,j)\in\Delta} y_{i,j} A_{i,j}.$$

$\bar{X}$ satisfies the desired property.

2.1.13. BFGS update Given a matrix C in $\mathbb{S}^n_{++}$ and vectors s and y in $\mathbb{R}^n$ satisfying $\langle s, y\rangle > 0$, consider the problem

$$\inf\{\langle C, X\rangle - \log\det(X) : Xs = y, X \in \mathbb{S}^n_{++}\}.$$

- Prove that for the problem above, the point

$$X = \frac{(y - \delta s)(y - \delta s)^T}{\langle s, y - \delta s\rangle} + \delta I$$

  is feasible for small $\delta > 0$.

- Prove problem has an optimal solution.

- Use *"First order conditions for linear constraints"* to find the solution. (The solution is called BFGS update of C^{-1} under the secant condition $Xs = y$.)

Proof:

- Note that

$$Xs = \frac{(y - \delta s)(y - \delta s)^T s}{\langle s, y - \delta s\rangle} + \delta X = y - \delta X + \delta X = y.$$

  Also, for $\delta > 0$ small enough, $\langle s, y - \delta s\rangle > 0$ and thus X will be the sum of a positive semi-definite matrix with δI and thus positive definite. Therefore, X is feasible for small $\delta > 0$.

- We know the map $X \in \mathbb{S}^n_{++} \mapsto \langle C, X\rangle - \log\det(X)$ has compact level sets and also $\{X \in \mathbb{S}^n : Xs = y\}$ is a closed affine subspace of $\mathbb{S}^n$. Thus the map $X \in \mathbb{S}^n_{++} \cap \{X \in \mathbb{S}^n : Xs = y\} \mapsto \langle C, X\rangle - \log\det(X)$ has compact level sets.

- From *"First order conditions for linear constraints"* for a local minimum $\bar{X}$ we know there exists w such that

$$C - \bar{X}^{-1} = sw^T + ws^T \Rightarrow (C - (sw^T + ws^T))^{-1} = \bar{X}.$$

  Therefore,

$$(C - (sw^T + ws^T))y = s \Rightarrow (sw^T + ws^T)y = Cy - s.$$

  So,

$$y^T(sw^T + ws^T)y = y^T Cy - y^T s \Rightarrow 2\langle y, s\rangle\langle y, w\rangle = y^T Cy - y^T s.$$

  Now,

$$\langle w, y\rangle = \frac{y^T Cy - \langle s, y\rangle}{2\langle s, y\rangle}.$$

2.1.15. Nearest polynomial with a given root
Consider the Euclidean space of complex polynomials of degree no more than n, with inner product

$$\left\langle \sum_{j=0}^{n} x_j z^j, \sum_{j=0}^{n} y_j z^j \right\rangle = \sum_{j=0}^{n} \overline{x_j} y_j.$$

Given a polynomial p in this space, calculate the nearest polynomial with a given complex root α, and prove the distance to this polynomial is

$$\left(\sum_{j=0}^{n} |\alpha|^{2j}\right)^{-\frac{1}{2}} ||p(\alpha)||.$$

Proof: The problem translates into the following optimization problem

$$\inf\{\sum_{i=0}^{n} ||a_i - b_i||^2 : \langle b, \overline{\tilde{\alpha}} \rangle = 0\},$$

where $p(x) = \sum_{j=0}^{n} a_i x^i$ and $\tilde{\alpha} = (1, \alpha, \cdots, \alpha^n)^T \in \mathbb{R}^{2n+2}$. Note that $\langle ., . \rangle : \mathbb{R}^{2n+2} \to \mathbb{R}^2$ is a linear map. Then for $b \in \mathbb{R}^{2n+2}$ a local minimizer to the above problem, we have

$$b - a = z\tilde{\alpha} \text{ for some } z \in \mathbb{C}$$

Now note that

$$b_i - a_i = z\alpha^i \Rightarrow 0 = \sum_{i=0}^{n} b_i \alpha^i = p(\alpha) + z \sum_{i=0}^{n} ||\alpha||^{2i}$$

So,

$$\sum_{i=0}^{n} ||a_i - b_i||^2 ||z||^2 \sum_{i=0}^{n} |\alpha|^{2i} \Rightarrow \sqrt{\sum_{i=0}^{n} ||a_i - b_i||^2} = \left(\sum_{j=0}^{n} |\alpha|^{2j}\right)^{-\frac{1}{2}} ||p(\alpha)||.$$

2 Max Functions

2.3.3
Prove by induction that if the functions $g_1, \cdots, g_m : \mathbb{E} \to \mathbb{R}$ are all continous at the point $\bar{x}$ then so is the max-function $g(x) = \max_i\{g_i(x)\}$.

Proof: It is clearly enough to prove the question for $m = 2$. If $g_1(\bar{x}) \neq g_2(\bar{x})$ then for instance if $g_1(\bar{x}) < g_2(\bar{x})$, then due to continuity for a small neighborhood about $\bar{x}$, g_1 is smaller than g_2 and thus g is equal to g_2. Since g_2 is continuous at $\bar{x}$ so is g.

So suppose that $g_1(\bar{x}) = g_2(\bar{x})$ and let $x^k \in \mathbb{E}$ be a converging sequence to $\bar{x}$. We aim to show $g(x^k) \to g(\bar{x})$. However, let R_1, R_2 be two subsequences of $\mathbb{N}$ such that for all $r \in R_i$, $g(x^r) = g_i(x^r)$, note that $R_1 \cap R_2$ is not necessarily empty, but $R_1 \cup R_2 = \mathbb{N}$. Now, suppose R_1 and R_2 are both infinite sized, then $g(x^k)$ is divided into two subsequences which both converge to $g(\bar{x}) = g_1(\bar{x}) = g_2(\bar{x})$. If only one of R_1 and R_2 are infinite, say for instance R_1, then $\lim g(x^k) = \lim g_1(x^k) = g_1(\bar{x})$ which equals to $g(\bar{x})$. The proof is complete.

2.3.5. Cauchy-Schwarz and steepest descent
For a nonzero vector y in $\mathbb{E}$, use Karush-Kuhn-Tucker conditions to solve the problem

$$\inf\{\langle y, x \rangle : ||x||^2 \leq 1\}$$

Deduce the Cauchy-Schwarz inequality.

Proof: Note that the feasible region is compact and the objective function is linear and thus continous. So there exists an optimal solution, not necessary unique, detoned by $\bar{x}$. Suppose $\bar{x} \neq 0$, then $\langle \nabla g(\bar{x}), -\bar{x} \rangle = 2\langle \bar{x}, -\bar{x} \rangle < 0$ and thus Mangasarian-Fromovitz constraint qualification holds at $\bar{x}$. Hence, there exists $\lambda \in \mathbb{R}_+$ such that

$$y + \lambda \bar{x} = 0.$$

So, since $y \neq 0$, λ is also nonzero and thus $||\bar{x}|| = 1$ and also $\bar{x} = -\frac{1}{\lambda} y$. Thus $\lambda = ||y||$. Hence, the objective value at $\bar{x}$ equals to

$$\langle y, \frac{-y}{||y||} \rangle = -||y|| < 0,$$

which is negative and thus the assumption that $\bar{x} \neq 0$ is justified. Finally we have for $x \neq 0$

$$\langle y, \frac{x}{||x||} \rangle \geq -||y|| \Rightarrow \langle y, x \rangle \geq -||y||||x||.$$

Intechanging x with $-x$ results in, which also holds for $x \neq 0$.

$$-||y||||x|| \leq \langle y, x \rangle \leq ||y||||x||.$$

2.3.7.
Consider a matrix $A \in \mathbb{S}^n_{++}$ and a real $b > 0$.

- Assuming the problem
$$\inf\{-\log\det X : \operatorname{Tr} AX \leq b, X \in \mathbb{S}^n_{++}\}$$
has a solution, find it.
- Repeat using the objective function $\operatorname{Tr}(X^{-1})$.
Proof:
- Note that $X \mapsto A^{\frac{1}{2}} X A^{\frac{1}{2}}$ is a homeomorphism and thus the following problem has a solution
$$\inf\{-\log\det(A^{\frac{1}{2}} X A^{\frac{1}{2}}) : \operatorname{Tr} A^{\frac{1}{2}} X A^{\frac{1}{2}} \leq b, A^{\frac{1}{2}} X A^{\frac{1}{2}} \in \mathbb{S}^n_{++}\}.$$
Thus without loss of generality suppose that $A = I$. Thus we know the problem
$$\inf\{-\log\det X : \operatorname{Tr} X \leq b, X \in \mathbb{S}^n_{++}\}$$
has a solution. Now, for any feasible matrix X, we have
$$\sqrt[1/n]{\det(X)} \leq \frac{1}{n}\operatorname{Tr}(X) \leq \frac{b}{n} \Rightarrow \det(X) \leq (\frac{b}{n})^{1/n} \Rightarrow -\log\det(X) \geq -\log(\frac{b}{n})^{1/n}.$$
However, in the above equation, equality happens if and only if $\lambda_1(X) = \cdots = \lambda_n(X)$ and also $\operatorname{Tr}(X) = b$. So the optimal solution equals to $\bar{X} = \frac{b}{n} I$.
Now for the original problem if X is the optimal solution $\bar{X}$, then $A^{\frac{1}{2}} X A^{\frac{1}{2}} = \frac{b}{n} I$ and so $\bar{X} = \frac{b}{n} A^{-1}$.
- Note that $\langle A, -A \rangle < 0$ and thus MFCQ holds at any $X \in \mathbb{S}^n_{++}$. Now note that if $\bar{X}$ is a local minimizer for the problem then since MFCQ holds at $\bar{X}$ there exists $\lambda \geq 0$ such that
$$-\bar{X}^{-1} + \lambda A = 0 \Rightarrow \bar{X}^{-1} = \lambda A$$
Note that λ can't be zero and so $\operatorname{Tr}(A\bar{X}) = b$. Hence, $\operatorname{Tr}(\frac{1}{\lambda} I) = b$ and so $\lambda = \frac{n}{b}$. Finally, $\bar{X} = \frac{b}{n} A^{-1}$.

2.3.8. Minimum volume ellipsoid

- For a $y \in \mathbb{R}^n$ and the function $g : \mathbb{S}^n \to \mathbb{R}$ defined by $g(X) = ||Xy||2$, prove $\nabla g(X) = Xyy^T + yy^T X$ for all the matrices X in $\mathbb{S}^n$.

- Consider a set $\{y^1, \cdots, y^m\} \subseteq \mathbb{R}^n$. Prove this set spans $\mathbb{R}^n$ if and only if the matrix $\sum_i y^i(y^i)^T$ is positive definite.
- Prove the problem

$$\begin{gathered}\inf -\log\det X \\ \text{subject to } ||Xy^i||^2 - 1 \le 0 \text{ for } i = 1, 2, \cdots, m \\ X \in \mathbb{S}^n_{++}\end{gathered}$$

  has an optimal solution.
- Show that the Mangasarian-Fromovitz constraint qualification holds at $\bar{X}$ by considering the direction $d = -\bar{X}$.
- Write down the KKT conditions that $\bar{X}$ must satisfy.
- When $\{y^1, \cdots, y^n\}$ is the standard basis of $\mathbb{R}^n$, the optimal solution of the problem in part (c) is $\bar{X} = I$. Find the corresponding Lagrange multiplier vector.

  Proof:

- Let $A = yy^T \in \mathbb{S}^n$, then $g(X) = \operatorname{Tr}(XAX)$ and thus

$$\lim_{t\to 0}\frac{1}{t}(g(X+tY) - g(X)) = \lim_{t\to 0}\frac{1}{t}(t\operatorname{Tr}(XAY) + \operatorname{Tr}(AXY) + t^2\operatorname{Tr}(YAY)) = \operatorname{Tr}((XA + AX)Y).$$

  Thus $\nabla g(X) = XA + AX$. Note that in the above equation we are using the fact that $\operatorname{Tr}(YAX) = \operatorname{Tr}(AXY)$.
- Clearly, $\sum_i y^i(y^i)^T \succeq 0$, also note that

$$x^T\left[\sum_i y^i(y^i)^T\right]x = \sum_i \langle x, y^i\rangle^2 \Rightarrow \left[\sum_i y^i(y^i)^T\right]x = 0 \iff \langle x, y^i\rangle = 0 \;\forall i$$

  So $\operatorname{Ker}(\sum_i y^i(y^i)^T) = 0$ if and only if it doesn't exist a vector x such that $\langle x, y^i\rangle = 0$ for all i and this holds if and only if the set $\{y^1, \cdots, y^m\} \subseteq \mathbb{R}^n$ spans $\mathbb{R}^n$.

  Now suppose the vector $y^1, \cdots, y^m$ span $\mathbb{R}^n$.
- Denote the feasible region of the above problem by Ω. Let $A = \sum_i y^i(y^i)^T$. Then as $y^1, \cdots, y^m$ span $\mathbb{R}^n$, we have $A \succ 0$. Also for $X \in \Omega$, we have $\langle X^2, A\rangle \le n$. Thus if for $X \in \Omega$, $-\log\det X \le c$ for some $c \in \mathbb{R}$, then

$$\langle A, X^2\rangle - \log\det X^2 \le n - \frac{1}{2}\log\det X \le n - \frac{1}{2}c.$$

  But we know that the level sets of $\langle C, X\rangle - \log\det X$ are compact for any $C \in \mathbb{S}^n_{++}$ from section 1.2, Question 14. However, $X \mapsto X^2$ is a homeomorphism from $\mathbb{S}^n_{++}$ to $\mathbb{S}^n_{++}$. Thus the set of $x \in \Omega$ which satisfies $-\log\det X \le c$ lie in a compact set. Thus the optimum is obtained.

  Now suppose that $\bar{X}$ is an optimal solution for the problem in part (c).
- Note that for all i

$$\langle \bar{X}y^i(y^i)^T + y^i(y^i)^T\bar{X}, -\bar{X}\rangle = -2\operatorname{Tr}(\bar{X}y^i(y_i)^T\bar{X}) = -2 < 0.$$

- The KKT conditions are as the followings

$$\begin{gathered}\bar{X} \in \mathbb{S}^n_{++} \\ ||\bar{X}y^i||^2 - 1 \le 0 \text{ for } i = 1, 2, \cdots, m \\ \exists\lambda \in \mathbb{R}^m_+,\ s.t.\ -\bar{X}^{-1} + \sum_i \lambda_i(\bar{X}y^i(y^i)^T + y^i(y^i)^T\bar{X}) = 0 \\ \lambda_i(||Xy^i||^2 - 1) = 0\;\forall i\end{gathered}$$

- Note that $I \in \Omega$ and $\log\det I = 0$. So we need to show that for $X \in \Omega$ we have $\log\det X \leq 0$ or equivalently $\det X \leq 1$. However, as we mentioned before, $\langle X, \sum_i y^i (y^i)^T \rangle \leq n$, but $\sum_i y^i (y^i)^T = I$ and thus $\langle X, I \rangle = \mathrm{Tr}(X) \leq n$. Now if $\lambda_1(X), \cdots, \lambda_n(X) > 0$ are the eigenvalues of X, the we have

$$\sqrt[1/n]{\prod_i \lambda_i(X)} \leq \frac{1}{n}(\sum_i \lambda_i(X)) = \frac{1}{n}\mathrm{Tr}(X) \leq 1 \Rightarrow \det(X) = \prod_i \lambda_i(X) \leq 1.$$

So I is the optimal solution. Now if $\lambda \in \mathbb{R}^n_+$ is a Lagrange multiplier then

$$-I + 2\,\mathrm{diag}(\lambda) = 0 \Rightarrow \lambda_i = \frac{1}{2}\ \forall i.$$