Additional Exercises: Convexity

1. Why a real symmetric matrix will always have real (as opposed to complex) eigenvalues?

2. Prove the following Cauchy-Schwarz inequality.

   For any $u, v \in \mathbf{R}^n$, we have
   $$u^{\mathrm{T}} v \leq \|u\|_2 \cdot \|v\|_2.$$

3. Use the Cauchy-Schwarz inequality to prove the so-called triangle inequality for the Euclidean norm:
   $$\|x+y\|_2 \leq \|x\|_2 + \|y\|_2$$
   for all $x, y \in \mathbf{R}^n$.

4. For a square matrix $A \in \mathbf{R}^{n\times n}$, its *trace* is $\operatorname{tr}(A) = \sum_{i=1}^n a_{ii}$. Prove: For any $X \in \mathbf{R}^{m\times n}$ and $Y \in \mathbf{R}^{n\times m}$, we have $\operatorname{tr}(XY^{\mathrm{T}}) = \operatorname{tr}(YX^{\mathrm{T}}) = \sum_{i=1}^m \sum_{j=1}^n X_{ij}Y_{ij}$.

5. Let $X \in \mathbf{R}^{m\times n}$ be a real matrix. The so-called Frobenius norm of X is defined as
   $$\|X\|_F := \left(\sum_{i=1}^m \sum_{j=1}^n X_{ij}^2 \right)^{1/2}$$
   and its spectrum norm is defined as $\|X\|_2 := \Big(\lambda_{\max}(X^{\mathrm{T}}X)\Big)^{1/2}$. Prove: Both $\|\cdot\|_F$ and $\|\cdot\|_2$ are indeed matrix norms.

6. Prove: For any $X \in \mathbf{R}^{m\times n}$ and $y \in \mathbf{R}^m$,
   $$\|Xy\|_2 \leq \|X\|_2 \cdot \|y\|_2.$$

7. Prove: For any X, it holds that $\|X\|_2 \leq \|X\|_F$.

8. Compute the gradient of the quartic function
   $$f(x) = (x^{\mathrm{T}}Ax)^2$$
   where $A \in \mathcal{S}^n$.

9. Compute the Hessian matrix of the quartic function
   $$f(x) = (x^{\mathrm{T}}Ax)^2$$
   where $A \in \mathcal{S}^n$.

10. Prove: If $h(x)$ is twice continuously differentiable, then $h(x)$ is convex in $\mathbf{R}^n$ is equivalent to $\nabla^2 h(x) \succeq 0$ for all $x \in \mathbf{R}^n$.

11. Prove: $(\prod_{i=1}^n x_i)^{1/n}$ is a concave function in $\mathbf{R}^n_{++}$.

12. Prove:
$$\frac{x_1^n}{x_2 x_3 \cdots x_n}$$
is a convex function in $\mathbf{R}^n_{++}$.

13. Consider $X \in \mathcal{S}^{n\times n}$, and so X has n real eigenvalues as we discussed before. Let them be
$$\lambda_1(X) \geq \lambda_2(X) \geq \cdots \geq \lambda_n(X).$$
Prove: $\lambda_1(X)$ is a convex function.

14. Prove:
$$\ln\left(\sum_{i=1}^n e^{x_i}\right)$$
is a convex function.

15. Suppose that $f(x) \geq 0$ is convex for $x \in S$, and $g(x) > 0$ is concave for $x \in S$. Prove:
$$\frac{f(x)}{g(x)}$$
is a quasi-convex function.

16. Show that
$$\frac{a^{\mathrm{T}}x + b}{c^{\mathrm{T}}x + d}$$
is quasi-linear in $\{x \mid c^{\mathrm{T}}x + d > 0\}$.

17. Suppose that $f(x) \geq 0$ is convex for $x \in S$, and $g(x) > 0$ is concave for $x \in S$. Prove:
$$\frac{f(x)^2}{g(x)}$$
is a convex function.

18. Prove: $\prod_{i=1}^n x_i$ is quasi-concave in $\mathbf{R}^n_{++}$.

19. Show that $S := \{x \mid \|x-a\|_2 \leq \|x-b\|_2\}$ is a convex region. Further prove: $\|x-a\|_2/\|x-b\|_2$ is quasi-convex in S.

20. Prove:
$$\Phi(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-t^2/2}dt$$
is a log-concave function.

21. Suppose $Q \in \mathcal{S}^{n\times n}_{++}$. Prove:

$$2x^{\mathrm{T}}y \le x^{\mathrm{T}}Qx + y^{\mathrm{T}}Q^{-1}y$$

for any $x, y \in \mathbf{R}^n$.

22. Suppose $0 < p < 1$. Show that

$$\left(\sum_{i=1}^{n} x_i^p\right)^{1/p}$$

is a *concave* function in $\mathbf{R}^n_{++}$.

23. If $f(x)$ is twice continuously differentiable and quasi-convex, then for any $x \in \mathrm{dom}(f)$:

$$d^{\mathrm{T}}\nabla f(x) = 0 \Longrightarrow d^{\mathrm{T}}\nabla^2 f(x) d \ge 0.$$

24. Prove: If the above condition holds, then there must exist some real value α such that

$$\nabla^2 f(x) + \alpha \nabla f(x)(\nabla f(x))^{\mathrm{T}} \succeq 0.$$

[The Hessian matrix of a quasi-convex function can have at most one negative eigenvalue!]

25. For $X \in \mathcal{S}^{n\times n}$, its eigenvalues are denoted to be

$$\lambda_1(X) \ge \lambda_2(X) \ge \cdots \ge \lambda_{n-1}(X) \ge \lambda_n(X).$$

Let $1 \le k \le n$. Consider

$$f(X) := \sum_{i=1}^{k} \lambda_i(X).$$

Prove: $f(X)$ is a convex function.

Hint: Show that

$$f(X) = \sup\{\mathrm{tr}(U^{\mathrm{T}}XU) \mid U \in \mathbf{R}^{n\times k},\, U^{\mathrm{T}}U = I_k\}.$$

26. A function $f:\ \mathbf{R}^n_{++} \to \mathbf{R}$

$$h(x) = c x_1^{\lambda_1} x_2^{\lambda_2} \cdots x_n^{\lambda_n}$$

with $c > 0$ and $\lambda \in \mathbf{R}^n$ is called a *monomial*. Sum of monomials, $f(x) = \sum_{i=1}^{k} h_i(x)$, is called a *posynomial*.

The so-called *geometric programming* problem is as follows:

$$\begin{array}{lll} (G) & \min & f_0(x) \\ & \text{s.t.} & f_i(x) \le 1,\ i = 1, 2, ..., m \\ & & h_j(x) = 1,\ j = 1, 2, ..., p \end{array}$$

where $f_i(x)$ are posynomials ($i = 1, 2, ..., m$), and $h_j(x)$ are monomials ($j = 1, 2, ..., p$).

Show that (G) can be formulated as convex optimization through a variable transformation.

27. Formulate the following L_4-norm approximation problem as QCQP:

$$\min \ \ \|Ax-b\|_4 = \left(\sum_{i=1}^{m}(a_i^{\mathrm{T}}x-b_i)^4\right)^{1/4}.$$

28. The so-called *Chebyshev center* of a polyhedron is the deepest point inside the polyhedron. Suppose that the polyhedron is given by $P=\{x \mid a_i^{\mathrm{T}}x \le b_i,\ i=1,2,...,m\}$. Formulate the problem of finding the Chebyshev center of P by a convex optimization model.

29. An ellipsoid may be given by the image of a ball under some linear transformation, e.g. $E = \{Bu+b \mid \|u\|_2 \le 1\}$. Without losing generality we can also assume $B \succ 0$. Then the volume of E is proportional to $\det B$.

    Consider again the polyhedron $P=\{x \mid a_i^{\mathrm{T}}x \le b_i,\ i=1,2,...,m\}$. Now the problem is to find the maximum volume ellipsoid inscribed inside P. Formulate the problem by convex optimization.

30. Let $A_i \in \mathcal{S}^{n\times n}$, $i=1,2,...,m$. Therefore, $A_0+x_1A_1+\cdots+x_mA_m$ is a symmetric matrix. We wish to find the values of $x_1,...,x_m$ so as to minimize the gap between the largest and the smallest eigenvalues of $A_0+x_1A_1+\cdots+x_mA_m$. Formulate this problem by SDP.

31. Let

$$\mathcal{K} := \{x \in \mathbf{R}^n \mid x_1 \ge x_2 \ge \cdots \ge x_n \ge 0\}.$$

    Show that $\mathcal{K}$ is a proper cone.

32. Find $A \in \mathbf{R}^{n\times n}$ such that $\mathcal{K} = A\,\mathbf{R}^n_+$.

33. In general, if $\mathcal{K} \subseteq \mathbf{R}^n$ is a proper cone, and $M \in \mathbf{R}^{n\times n}$ is a non-singular matrix, then $M\mathcal{K}$ is also a proper cone.

34. Compute $(M\mathcal{K})^*$.

35. Derive the dual of the following *non-standard* conic optimization problem:

$$\begin{array}{ll} \min & c^{\mathrm{T}}x \\ \text{s.t.} & A_1x+b_1 \in \mathcal{K}_1 \\ & A_2x+b_2 \in \mathcal{K}_2 \\ & \vdots \\ & A_mx+b_m \in \mathcal{K}_m, \end{array}$$

    where $\mathcal{K}_1, \mathcal{K}_2, \cdots, \mathcal{K}_m$ are all closed convex cones.

36. Suppose that $f(x)$ is a convex function, and its conjugate function is known to be $f^*(s)$. Consider the following optimization model

$$\begin{array}{ll} \min & f(x) \\ \text{s.t.} & Ax \leq b. \end{array}$$

Derive the Lagrangian dual of the above problem.

37. The channel capacity optimization problem is:

$$\begin{array}{ll} \min & -c^{\mathrm{T}}x + \sum_{i=1}^{m} y_i \ln y_i \\ \text{s.t.} & Px = y \\ & x \geq 0,\ \mathbf{1}^{\mathrm{T}}x = 1. \end{array}$$

What is the dual of the above problem?

38. The sum of first k largest components of vector $x \in \mathbf{R}^n$ ($k < n$) is known to be a convex function. (Why?) Denote this function to be $f(x)$. Formulate the following portfolio selection problem using $f(x)$: We wish to select from a total of n assets to form a portfolio (no short-selling is allowed). Asset i has an expected rate of return $\mu_i > 0$, and the covariance matrix is Σ. We wish to minimize the variance of the portfolio while requiring that the expected rate of return to the portfolio is at least μ. Moreover, the weight of the first k largest components of investment should not exceed half of the total investment.

39. The condition that $f(x) \leq 0.5$ can be formulated by linear programming. How?

Solutions (Convexity)

1. Let $Ax = \lambda x$. Then, $x^H A x = \lambda x^H x$

$$\Rightarrow \overline{(x^H A x)} = \bar{\lambda} \cdot x^H x = x^H A^H x = x^H A x = \lambda x^H x$$

$$\Rightarrow \bar{\lambda} = \lambda \; : \text{ a real value.}$$

2. $\forall$ real value t : $0 \le \| u + t \cdot v \|_2^2 = \|u\|_2^2 + 2t \cdot u^T v + t^2 \cdot \|v\|_2^2$

$\Rightarrow$ Its discreminant $\Delta = (u^T v)^2 - \|u\|_2^2 \cdot \|v\|_2^2 \le 0$

$$\Rightarrow |u^T v| \le \|u\|_2 \cdot \|v\|_2 .$$

3.

$$\|x+y\|_2^2 = \|x\|_2^2 + 2x^T y + \|y\|_2^2 \le \|x\|_2^2 + 2\|x\|_2 \cdot \|y\|_2 + \|y\|_2^2$$
$$= (\|x\|_2 + \|y\|_2)^2$$

$$\Rightarrow \|x+y\|_2 \le \|x\|_2 + \|y\|_2 .$$

4.

$$\mathrm{tr}(XY^T) = \sum_{j=1}^{m} \left(\sum_{i=1}^{n} X_{ji} \cdot Y_{ji} \right) = \sum_{i=1}^{n} \left(\sum_{j=1}^{m} Y_{ij} X_{ij} \right) = \mathrm{tr}(YX^T)$$

5. Let the SVD (Singular Value Decomposition) of X be

$$X = U \cdot \Sigma \cdot V^T = \sum_{i=1}^{r} \sigma_i u_i \cdot v_i^T$$

where $u_i \in \mathbb{R}^m$, $v_j \in \mathbb{R}^n$, and they are orthonormal in their own domain.

We have $\|X\|_F^2 = \mathrm{tr}(XX^T) = \sum_{i=1}^{r} \sigma_i^2$ and $\|X\|_2 = \max_{1 \le i \le r} \sigma_i$.

Let us verify that the triangle inequality is satisfied by these matrix norms. (The other two definitions of a norm, i.e., $\|x\| = 0 \iff x = 0$; $\|tx\| = |t| \cdot \|x\|$, are trivial to verify).

That $\|X\|_F$ is a norm follows from the fact that $\|X\|_F$ is the Euclidean norm on the matrix X as a vector.

To verify that $\|X\|_2$ is a matrix norm, we proceed to the next question first.

$\boxed{6.}$ Suppose $X^T X v = \lambda_{max} \cdot v$ where λ_{max} is the eigenvalue with largest absolute value. Then, $v^T X^T X v = \lambda_{max} v^T v$.

On the other hand, $y^T X^T X y \le \lambda_{max} \cdot y^T y = \|X\|_2^2 \cdot y^T y$.

Therefore, $\max\limits_{\|y\|_2 = 1} \|Xy\|_2 = \|X\|_2$.

Going back to Question 5 : $\|X + Y\|_2 = \max\limits_{\|v\|_2 = 1} \|(X+Y) \cdot v\|_2$

$\le \max\limits_{\|v\|_2 = 1} \|X \cdot v\|_2 + \max\limits_{\|v\|_2 = 1} \|Yv\|_2 = \|X\|_2 + \|Y\|_2$.

The triangle inequality is shown.

7. $$\|X\|_2 = \max_{1\le i\le r} \sigma_i \le \left(\sum_{i=1}^{r} \sigma_i^2\right)^{\frac{1}{2}} = \|X\|_F .$$

8. $$\nabla\left((x^TAx)^2\right) = 2\cdot(x^TAx)\cdot \nabla(x^TAx) = 4\cdot(x^TAx)\cdot Ax .$$

9. $$\nabla^2\left((x^TAx)^2\right) = 8\, Ax\cdot(Ax)^T + 4\cdot(x^TAx)\cdot A .$$

10. $h(x)$ is convex & smooth $\Longleftrightarrow$ $h(y) \ge h(x) + \nabla h(x)^T(y-x)$ $\forall\, x, y$

If $\nabla^2 h(z) \succeq 0$ $\forall\, z$, then by the Taylor expression

$$h(y) - h(x) - \nabla h(x)^T(y-x) = \frac{1}{2}(y-x)^T\cdot \nabla^2 f(z)\cdot(y-x) \succeq 0$$

where z is between x and y.

On the other hand, if $h(y) \ge h(x) + \nabla h(x)^T(y-x)$ $\forall\, x, y$, then one may choose $y = x + \Delta x$ to derive

$$\Delta x^T\cdot \nabla^2 f(z)\cdot \Delta x \ge 0$$

Because Δx can be chosen arbitrarily, we have $\nabla^2 f(x) \succeq 0$.

11. Let $f(x) = \left(\prod_{i=1}^{n} x_i\right)^{\frac{1}{n}} = e^{\frac{1}{n}\left(\sum_{i=1}^{n} \ln x_i\right)}$

$$\nabla f(x) = e^{\frac{1}{n}\left(\sum_{i=1}^{n} \ln x_i\right)} \cdot \begin{pmatrix} \frac{1}{nx_1} \\ \vdots \\ \frac{1}{nx_n} \end{pmatrix}$$

$$\nabla^2 f(x) = e^{\frac{1}{n}\left(\sum_{i=1}^{n} \ln x_i\right)} \left[\frac{1}{n^2} \begin{pmatrix} \frac{1}{x_1} \\ \vdots \\ \frac{1}{x_n} \end{pmatrix} \left(\frac{1}{x_1} \cdots \frac{1}{x_n}\right) - \frac{1}{n} \begin{pmatrix} \frac{1}{x_1^2} & & 0 \\ & \ddots & \\ 0 & & \frac{1}{x_n^2} \end{pmatrix} \right]$$

$$\xi^T \nabla^2 f(x)\, \xi = e^{\frac{1}{n}\left(\sum_{i=1}^{n} \ln x_i\right)} \cdot \left[\left(\frac{1}{n}\sum_{i=1}^{n} \frac{\xi_i}{x_i}\right)^2 - \frac{1}{n}\sum_{i=1}^{n} \frac{\xi_i^2}{x_i^2} \right] \le 0$$

Cauchy-Schwarz!

$\Longrightarrow$ f is concave.

$\boxed{12}$. Let $f(x) = \dfrac{x_1^n}{x_2 \cdots x_n}$

$$\ln f(x) = n \cdot \ln x_1 - \sum_{i=2}^{n} \ln x_i =: g(x)$$

$$\nabla g(x) = \begin{pmatrix} \frac{n}{x_1} \\ -\frac{1}{x_2} \\ \vdots \\ -\frac{1}{x_n} \end{pmatrix}, \quad \nabla^2 g(x) = \begin{pmatrix} -\frac{n}{x_1^2} & & & 0 \\ & \frac{1}{x_2^2} & & \\ & 0 & \ddots & \\ & & & \frac{1}{x_n^2} \end{pmatrix}$$

Because $f(x) = e^{g(x)}$,

$$\nabla f(x) = f(x) \cdot \nabla g(x), \quad \nabla^2 f(x) = f(x) \cdot \nabla^2 g(x) + f(x) \cdot \nabla g(x) \cdot \nabla g(x)^T$$

$$\Longrightarrow \nabla^2 f(x) = f(x) \cdot \left[\begin{pmatrix} -\frac{n}{x_1^2} & & & \\ & \frac{1}{x_2^2} & & \\ & & \ddots & \\ & & & \frac{1}{x_n^2} \end{pmatrix} + \begin{pmatrix} \frac{n}{x_1} \\ -\frac{1}{x_2} \\ \vdots \\ -\frac{1}{x_n} \end{pmatrix} \left(\frac{n}{x_1} \;\; -\frac{1}{x_2} \;\; \cdots \;\; -\frac{1}{x_n} \right) \right]$$

Take any $\xi \in \mathbb{R}^n$:

$$\xi^T \nabla^2 f(x) \xi = f(x) \cdot \left[-n \cdot \left(\frac{\xi_1}{x_1}\right)^2 + \sum_{i=2}^{n} \left(\frac{\xi_i}{x_i}\right)^2 + \left(n \cdot \frac{\xi_1}{x_1} - \sum_{i=2}^{n} \frac{\xi_i}{x_i}\right)^2 \right].$$

To show that the above expression is always non-negative, denote $z_i = \frac{\xi_i}{x_i}$, $i = 1, \cdots, n$. We have

$$-n z_1^2 + \sum_{i=2}^{n} z_i^2 + \left(n z_1 - \sum_{i=2}^{n} z_i\right)^2$$

$$= (z_1, \cdots, z_n) \left[\begin{pmatrix} -n & & \\ & 1 & \\ & & \ddots & \\ & & & 1 \end{pmatrix} + \begin{pmatrix} n \\ -1 \\ \vdots \\ -1 \end{pmatrix} (n, -1, \cdots, -1) \right] \begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix} =: z^T Q \cdot z.$$

We write Q as a block matrix

$$Q = \begin{pmatrix} -n & \\ & I_{n-1} \end{pmatrix} + \begin{pmatrix} n \\ -\mathbb{1}_{n-1} \end{pmatrix} (n, -\mathbb{1}_{n-1}^T)$$

$$= \begin{pmatrix} n^2 - n & -n \cdot \mathbb{1}_{n-1}^T \\ -n \cdot \mathbb{1}_{n-1} & I_{n-1} + \mathbb{1}_{n-1} \cdot \mathbb{1}_{n-1}^T \end{pmatrix}$$

By the so-call Schur Complement Lemma: $Q \succeq 0$

$$\Longleftrightarrow \quad I_{n-1} + \mathbb{1}_{n-1} \cdot \mathbb{1}_{n-1}^T - \frac{n^2}{n^2-n} \cdot \mathbb{1}_{n-1} \cdot \mathbb{1}_{n-1}^T = I_{n-1} - \frac{1}{n-1} \mathbb{1}_{n-1} \cdot \mathbb{1}_{n-1}^T \succeq 0$$

But the above matrix inequality is obvious, because $\mathbb{1}_{n-1} \cdot \mathbb{1}_{n-1}^T$ has eigenvalue 0 with multiplicity $n-2$ and eigenvalue $n-1$. Therefore, $\mathbb{1}_{n-1} \cdot \mathbb{1}_{n-1}^T \preceq (n-1) \cdot I_{n-1}$.

$\boxed{13.}$ We use the fact that if $f(x;a)$ is a convex function in x for any fixed a, then $\max\limits_{a \in A} f(x;a)$ remains a convex function in x.

Because $\lambda_1(X) = \max\limits_{\|v\|_2 = 1} \dfrac{v^T X v}{\|v\|_2^2}$ and $\dfrac{v^T X v}{\|v\|_2^2}$ is a linear function in X for fixed v, so $\lambda_1(X)$ is a convex function in X.

$\boxed{14.}$ Let $f(x) = \ln\left(\sum_{i=1}^{n} e^{x_i}\right)$.

$$\nabla f(x) = \frac{1}{\sum_{i=1}^{n} e^{x_i}} \begin{pmatrix} e^{x_1} \\ \vdots \\ e^{x_n} \end{pmatrix} \quad \text{and}$$

$$\nabla^2 f(x) = \frac{1}{\sum_{i=1}^{n} e^{x_i}} \begin{pmatrix} e^{x_1} & & 0 \\ & \ddots & \\ 0 & & e^{x_n} \end{pmatrix} - \frac{1}{\left(\sum_{i=1}^{n} e^{x_i}\right)^2} \begin{pmatrix} e^{x_1} \\ \vdots \\ e^{x_n} \end{pmatrix} \left(e^{x_1} \cdots e^{x_n}\right)$$

Take any $\xi \in \mathbb{R}^n$:

$$\xi^T \nabla^2 f(x)\, \xi = \frac{1}{\sum_{i=1}^{n} e^{x_i}} \left(\sum_{i=1}^{n} \xi_i^2 \cdot e^{x_i}\right) - \frac{1}{\left(\sum_{i=1}^{n} e^{x_i}\right)^2} \left(\sum_{i=1}^{n} \xi_i e^{x_i}\right)^2$$

$$= \frac{1}{\left(\sum_{i=1}^{n} e^{x_i}\right)^2} \left[\left(\sum_{i=1}^{n} e^{x_i}\right)\left(\sum_{i=1}^{n} \xi_i^2 e^{x_i}\right) - \left(\sum_{i=1}^{n} \xi_i e^{x_i}\right)^2 \right] \geq 0$$

where we used the Cauchy-Schwarz inequality :

[let $a_i := \xi_i e^{\frac{x_i}{2}}$, $b_i := e^{\frac{x_i}{2}}$

then $\left(\sum_{i=1}^{n} a_i b_i\right)^2 \leq \left(\sum_{i=1}^{n} a_i^2\right) \cdot \left(\sum_{i=1}^{n} b_i^2\right)$.]

The convexity of $f(x)$ follows, since $\nabla^2 f(x) \succeq 0$ as we proved above.

15. The level set of $\frac{f}{g}$ is:

$$L_\lambda\left(\frac{f}{g}\right) = \left\{ x \,\middle|\, \frac{f(x)}{g(x)} \le \lambda \right\} = \{ x \mid f(x) - \lambda \cdot g(x) \le 0 \}$$

Since $f(x) - \lambda g(x)$ is convex, we know that $L_\lambda\left(\frac{f}{g}\right)$ is a convex set. Hence, $\frac{f}{g}$ is quasi-convex.

16. $\dfrac{a^T x + b}{c^T x + d}$ is both quasi-convex and quasi-concave (hence quasi-linear) because a linear function is both convex and concave.

17. It is easy to verify that: if $F(x_1, x_2)$ is convex, and $F(x_1, x_2)$ is increasing in x_1 for fixed x_2, and $F(x_1, x_2)$ is decreasing in x_2 for fixed x_1, then $F(f(x), g(x))$ is convex if f is convex and $g(x)$ is concave. This is because

$$f(\lambda x + (1-\lambda) y) \le \lambda f(x) + (1-\lambda) f(y) \quad \& \quad g(\lambda x + (1-\lambda) y) \ge \lambda g(x) + (1-\lambda) g(y)$$

$$\Longrightarrow F(f(\lambda x + (1-\lambda) y), g(\lambda x + (1-\lambda) y)) \overset{\text{monotonicity}}{\le} F(\lambda f(x) + (1-\lambda) f(y), \lambda g(x) + (1-\lambda) g(y))$$

$$\overset{\text{convexity of } F}{\le} \lambda F(f(x), g(x)) + (1-\lambda) F(f(y), g(y)).$$

Now we use what we proved in Question 12:

$$\frac{x_1^2}{x_2} \quad \text{is convex in } (x_1, x_2) \quad \text{for } (x_1, x_2) > 0.$$

We have $\dfrac{f(x)^2}{g(x)}$ is a convex function.

$\boxed{18.}$ $\displaystyle\prod_{i=1}^{n} x_i = e^{\sum_{i=1}^{n} \ln x_i}$.

$$\Rightarrow \left\{ x \,\middle|\, x > 0,\ \prod_{i=1}^{n} x_i \geq t \right\} = \left\{ x \,\middle|\, x>0,\ \sum_{i=1}^{n} \ln x_i \geq \ln t \right\}$$

which is a convex set for all $t > 0$. Therefore, $\prod_{i=1}^{n} x_i$ is quasi-concave for $x > 0$.

$\boxed{19.}$ $$\left\{ x \,\middle|\, \|x-a\|_2 \leq \|x-b\|_2 \right\} = \left\{ x \,\middle|\, x^T x - 2a^T x + a^T a \leq x^T x - 2b^T x + b^T b \right\}$$

$$= \left\{ x \,\middle|\, 2(b-a)^T x \leq b^T b - a^T a \right\}$$ which is a half-space, hence a convex set.

$\boxed{20.}$ We want to show: $\ln\left(\int_{-\infty}^{x} e^{-t^2/2}\,dt\right) =: f(x)$

is a concave function.

We have: $f'(x) = \dfrac{e^{-x^2/2}}{\int_{-\infty}^{x} e^{-t^2/2}\,dt}$

$$\Rightarrow \quad f''(x) = -\frac{e^{-\frac{x^2}{2}}\cdot x}{\int_{-\infty}^{x} e^{-t^2/2}\,dt} - \frac{\left(e^{-x^2/2}\right)^2}{\left(\int_{-\infty}^{x} e^{-t^2/2}\,dt\right)^2} < 0.$$

$\boxed{21.}$ For any positive definite matrix $Q \succ 0$, there is a unique positive matrix $Q^{\frac{1}{2}}$ such that $Q^{\frac{1}{2}}\cdot Q^{\frac{1}{2}} = Q$ and $Q^{\frac{1}{2}} \succ 0$.

Therefore, $0 \le \| Q^{\frac{1}{2}}x - Q^{-\frac{1}{2}}y \|_2^2 \le x^TQx - 2x^TQ^{\frac{1}{2}}\cdot Q^{-\frac{1}{2}}y + y^TQ^{-1}y$

$$= x^TQx + y^TQ^{-1}y - 2x^Ty.$$

$\boxed{22.}$ Let $f(x) = \left(\sum_{i=1}^{n} x_i^p\right)^{\frac{1}{p}}$. We have

$$\nabla f(x) = \frac{1}{p}\left(\sum_{i=1}^{n} x_i^p\right)^{\frac{1}{p}-1}\cdot\begin{pmatrix} p\cdot x_1^{p-1} \\ \vdots \\ p\cdot x_n^{p-1}\end{pmatrix} = \left(\sum_{i=1}^{n} x_i^p\right)^{\frac{1-p}{p}}\cdot\begin{pmatrix} x_1^{p-1} \\ \vdots \\ x_n^{p-1}\end{pmatrix}$$

$$\nabla^2 f(x) = \frac{1-p}{p}\cdot\left(\sum_{i=1}^{n} x_i^p\right)^{\frac{1-2p}{p}}\cdot p\cdot\begin{pmatrix} x_1^{p-1} \\ \vdots \\ x_n^{p-1}\end{pmatrix}(x_1^{p-1}, \cdots, x_n^{p-1}) + \left(\sum_{i=1}^{n} x_i^p\right)^{\frac{1-p}{p}}\cdot(p-1)\cdot\begin{pmatrix} x_1^{p-2} & & 0 \\ & \ddots & \\ 0 & & x_n^{p-2}\end{pmatrix}$$

Take any $\xi \in \mathbb{R}^n$:

$$\xi^T \nabla^2 f(x)\, \xi = (p-1)\cdot\left(\sum_{i=1}^n x_i^p\right)^{\frac{1-2p}{p}} \left[-\left(\sum_{i=1}^n \xi_i \cdot x_i^{p-1}\right)^2 + \left(\sum_{i=1}^n x_i^p\right)\cdot\left(\sum_{i=1}^n \xi_i^2 \cdot x_i^{p-2}\right) \right]$$

Denote $a_i = x_i^{\frac{p}{2}}$, $b_i = \xi_i \cdot x_i^{\frac{p-2}{2}}$. We have

$$-\left(\sum_{i=1}^n \xi_i x_i^{p-1}\right)^2 + \left(\sum_{i=1}^n x_i^p\right)\cdot\left(\sum_{i=1}^n \xi_i^2 x_i^{p-2}\right) = -\left(\sum_{i=1}^n a_i b_i\right)^2 + \left(\sum_{i=1}^n a_i^2\right)\cdot\left(\sum_{i=1}^n b_i^2\right) \geq 0$$

Cauchy-Schwarz inequality!

Therefore $\xi^T \nabla^2 f(x)\, \xi \leq 0$ (notice $p < 1$) $\forall \xi$, and so f is concave.

$\boxed{23.}$ Suppose, by contradiction, that $d^T \nabla f(x) = 0$ and $d^T \nabla^2 f(x) d < 0$.

Consider $h(t) = f(x + td)$. We have $h'(t) = \nabla f(x+td)^T d$ and $h''(t) = d^T \nabla^2 f(x+td) d$.

In this case, $h'(0) = 0$ and $h''(0) < 0$. Therefore 0 is a local maximum for $h(t)$, for $t \in (-\varepsilon, \varepsilon)$ $(\varepsilon > 0)$.

So, there is $t > 0$ such that $h(t) < h(0)$ & $h(-t) < h(0)$

Then, $f(x) > \max\{f(x+td), f(x-td)\}$,

which is a contradiction to the definition of quasi-convexity.

24. This is basically a linear algebra exercise.

It says that if there is a real-symmetric matrix Q and a real-vector v, in such a way that

$$\forall d: d^T v = 0 \implies d^T Q d \geq 0$$

then there must exist a value t such that $Q + t v v^T \succeq 0$.

Let ~~use~~ us first prove this fact in a slighty stronger form:

$$\forall d \neq 0: d^T v = 0 \implies d^T Q d > 0$$

then there exists t such that $Q + t v v^T \succeq 0$.

To see this, consider $\displaystyle \tau := \max_{\substack{v^T d = 0 \\ d \neq 0}} \frac{(v^T Q d)^2}{d^T Q d} \geq 0$.

Choose t such that $t \cancel{v^T} v^T v + v^T Q v \geq \tau$.

Since any $x \in \mathbb{R}^n$ can be written as $x = \alpha \cdot v + \beta d$ with $v^T d = 0$, we have

$$x^T (Q + t \cdot v v^T) x = (\alpha v + \beta d)^T (Q + t v v^T)(\alpha v + \beta d)$$

$$= \alpha^2 (v^T Q v + t \cdot v^T v) + 2\alpha\beta \cdot v^T Q d + \beta^2 d^T Q d$$

$$\geq \tau \cdot \alpha^2 + 2\alpha\beta \cdot v^T Q d + \beta^2 \cdot d^T Q d$$

The discreminant of the above quadratic form is

$$\Delta = (v^T Q d)^2 - \tau \cdot d^T Q d \leq 0.$$

Therefore, $\tau \cdot \alpha^2 + 2\alpha\beta\, v^T Q d + \beta^2 d^T Q d \geq 0$

24 (Continued) Therefore $Q + t \cdot vv^T \succeq 0$.

We may choose Q to be $Q + \varepsilon \cdot I$ with $\varepsilon > 0$.

Then, $Q + \varepsilon \cdot I + t_\varepsilon \cdot v \cdot v^T \succeq 0 \quad \forall \varepsilon > 0$.

That is, the smallest eigenvalue of $Q + t_\varepsilon \cdot v \cdot v^T$ is at most $-\varepsilon$. This also implies that Q has only one negative eigenvalue with eigenvector v. Hence, there is ~~so~~ a value t^*, in such a way that $\forall\, t \geq t^*$ we have $Q + t\, v \cdot v^T \succeq 0$.

$\boxed{25.}$ Let $X = Q^T \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} Q$ where $Q^T Q = I$.

Take any $U \in \mathbb{R}^{n \times k}$ with $U^T \cdot U = I_k$. Let $\tilde{U} = QU$.

We have $\tilde{U}^T \tilde{U} = I_k$, and so

$$\mathrm{tr}(U^T X U) = \mathrm{tr}\left(\tilde{U}^T \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} \tilde{U}\right) = \mathrm{tr}\left[\begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} \cdot \tilde{U} \cdot \tilde{U}^T\right]$$

$= \sum_{i=1}^{n} \lambda_i \cdot u_i^2$, where u_i is the norm of the ith row of $\tilde{U}$. We have that $u_i^2 \leq 1 \; \forall i$ and $\sum_{i=1}^{n} u_i^2 = k$.

Therefore, $\mathrm{tr}(U^T X U) = \sum_{i=1}^{n} \lambda_i u_i^2 \leq \lambda_1 + \lambda_2 + \cdots + \lambda_k$.

On the other hand, we can choose U to be the k eigenvectors of X, and then $\mathrm{tr}(U^T X U) = \lambda_1 + \cdots + \lambda_k$.

25 (Continued.) Therefore, we have shown

$$f(X) = \max \{ \operatorname{tr}(U^T X U) \mid U \in \mathbb{R}^{n \times k},\ U^T U = I_k \}.$$

For any fixed U, $\operatorname{tr}(U^T X U)$ is a linear function in X, hence convex. Therefore, $f(X)$ is a convex function in X.

$\boxed{26.}$ Let us introduce a variable transformation

$$x_i := e^{y_i}, \quad i = 1, \cdots, m.$$

The monomial $h(x) = c \cdot x_1^{\lambda_1} \cdots x_n^{\lambda_n} = c \cdot e^{\lambda_1 y_1 + \cdots + \lambda_n y_n}$.

The constraint $h_j(x) = \phi$ becomes $\lambda_{j1} \cdot y_1 + \cdots + \lambda_{jn} y_n = c_j$

and the constraint $f_i(x) \le 1$ becomes

$$\sum_i c_i e^{\sum_{j=1}^{n} \lambda_{ij} y_j} \le 1 \implies \ln\left(\sum_i c_i e^{\sum_{j=1}^{n} \lambda_{ij} y_j} \right) \le 0.$$

According to Exercise 14, we know that

$$\ln\left(\sum_i c_i e^{\sum_{j=1}^{n} \lambda_{ij} y_j} \right)$$

is convex. Therefore, after the above transformation, the geometric programming problem becomes convex programming!

27.

$$\begin{aligned} \min \quad & \sum_{i=1}^{m} t_i \\ \text{s.t.} \quad & t_i \geq s_i^2 \\ & s_i \geq (a_i^T x - b_i)^2, \quad i=1,\cdots,m. \end{aligned}$$

28. Let the point be located at p, and the entire Euclidean ball with p as the center, radius t, is within the polyhedron. That is,

$a_i^T(p+t\xi) \leq b_i$, $i=1,\cdots,m$, and ξ is a unit vector. This is equivalent to: $a_i^T p + t\|a_i\| \leq b_i$, $i=1,\cdots,m$.

Therefore, the problem of finding the chebyshev center is:

$$\begin{aligned} \max \quad & t \\ \text{s.t.} \quad & a_i^T p + t \cdot \|a_i\| \leq b_i, \quad i=1,\cdots,m \end{aligned}$$

where p and t are the decision variables, and the problem is convex optimization.

29. Similar as in Exercise 28, the constraints are:

$a_i^T(Bu+b) \leq b_i$, $\forall \|u\| \leq 1$, which is: $\|Ba_i\| + a_i^T b \leq b_i$.

The problem is:

$$\begin{aligned} \max \quad & \ln(\det(B)) \\ \text{s.t.} \quad & \|Ba_i\| + a_i^T b \leq b_i, \quad i=1,\cdots,m \\ & B \succ 0 \end{aligned}$$

where the decision variables are: B, b.

30.

$$\min \quad z - y$$
$$\text{s.t.} \quad y \cdot I \preceq A_0 + x_1 A_1 + \cdots + x_m A_m \preceq z \cdot I.$$

31.

K is : (1) a convex cone (easy to verify)

(2) K is pointed : If $x \in K$ & $-x \in K$

$\Longrightarrow x = 0$

(3) K is solid :

$$\begin{pmatrix} n \\ n-1 \\ \vdots \\ 1 \end{pmatrix} + B \subseteq K,$$ where B is a unit ball.

32.

$$A = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ 0 & 1 & \cdots & 1 \\ 0 & 0 & 1 \cdots & 1 \\ \cdots & & \cdots & \\ 0 & 0 & 0 \cdots & 1 \end{pmatrix}$$

33.

(1) MK is a convex cone;

(2) If $Mx \in MK$ and $-Mx \in MK$

$\Rightarrow x$ & $-x \in K \Rightarrow x = 0 \Rightarrow M \cdot 0 = 0$

(3) A solid ball under non-singular linear Transformation becomes a solid ellipsoid.

34.

$$(M \cdot K)^* = \{ y \mid x^T M^T y \geq 0 \quad \forall x \in K \}$$
$$= \{ y \mid M^T y \in K^* \}$$
$$= M^{-T} \cdot K^*$$

35.

$$\max \quad \sum_{i=1}^{m} b_i^T y_i$$
$$\text{s.t.} \quad A_1^T y_1 + \cdots + A_m^T y_m + c = 0$$
$$y_i \in K_i^*, \quad i = 1, \cdots, m.$$

36.

$$\max \quad -b^T y - f^*(-A^T y)$$
$$\text{s.t.} \quad y \geq 0$$

$\boxed{37.}$

$$\max \quad -\sum_{i=1}^{m} e^{\lambda_i - 1} - \lambda_0$$

$$\text{s.t.} \quad -c + P^T \lambda + \lambda_0 \mathbb{1} \geq 0 .$$

$\boxed{38.}$

$$\min \quad x^T \Sigma x$$

$$\text{s.t.} \quad \mathbb{1}^T x = 1, \quad d^T x \geq \mu$$

$$f(x) \leq 1/2, \quad x \geq 0$$

where $\quad d^T = (\mu_1, \mu_2, \cdots, \mu_n)$.

$\boxed{39.}$ The constraint $f(x) \leq 1/2$ can be represented by:

$$x_{i_1} + x_{i_2} + \cdots + x_{i_k} \leq 1/2$$

for all $1 \leq i_1 < i_2 < i_3 < \cdots < i_k \leq n$, which is certainly polyhedral, but it involves $\binom{n}{k} = \frac{n!}{k!\,(n-k)!}$ (linear constraints.