ELEG5481 Signal Processing Optimization Techniques
Assignment Solution 3

Feb. 27, 2013

Q1. The following functions are either convex, concave, or neither convex nor concave. Identify their convexity/concavity, and provide your answer with a proof.

(a)

$$f(x) = \max\{\|APx - b\| \ \mid \ P \text{ is a permutation matrix } \}$$

with $A \in \mathbf{R}^{m \times n}$, $b \in \mathbf{R}^m$. Note that a permutation matrix P is a square matrix that has exactly one entry 1 in each row and column and 0s elsewhere.

(b)

$$f(x) = \|Ax - b\|_2^2 - \gamma \|x\|_2^2$$

where $\mathbf{dom} f = \mathbf{R}^n$, $A \in \mathbf{R}^{m \times n}$, $b \in \mathbf{R}^m$, and $\gamma > 0$.

(c)

$$f(X) = \lambda_{\min}(X) = \inf_{y \in \mathbf{R}^n, y \neq 0} \frac{y^T X y}{y^T y}, \quad \mathbf{dom} f = \mathbf{S}^n$$

(d)

$$f(x) = \int_0^{2\pi} \log p(x, \omega) d\omega,$$

where

$$p(x, \omega) = x_1 + x_2 \cos(\omega) + x_3 \cos(2\omega) + \ldots x_n \cos((n-1)\omega),$$

and $\mathbf{dom} f = \{x \mid p(x, \omega) > 0, 0 \leq \omega \leq 2\pi\}$. (Note that $\log(\cdot)$ is the natural log function.)

(e) The difference between the maximum and minimum value of a polynomial on a given interval, as a function of its coefficients:

$$f(x) = \sup_{t \in [a,b]} p(t) - \inf_{t \in [a,b]} p(t),$$

where $p(t) = x_1 + x_2 t + x_3 t^2 + \ldots + x_n t^{n-1}$, and a and b are real constants with $a < b$.

(f)

$$f(x) = \sum_{i=1}^{m} e^{-1/f_i(x)}, \quad \mathbf{dom} f = \{x \ \mid f_i(x) < 0, \ i = 1, \ldots, m\}.$$

where the functions f_i are convex.

Solution:

(a) For a given P, $\|APx - b\|$ is a convex function of x. Therefore

$$f(x) = \max_{P: P \text{ is a permuation matrx}} \|APx - b\|$$

is convex.

(b)

$$\begin{aligned} f(x) &= x^T A^T A x - 2x^T A^T b + \|b\|_2^2 - \gamma x^T x^T \\ &= x^T (A^T A - \gamma I) x - 2x^T A^T b + \|b\|_2^2 \end{aligned}$$

This is a quadratic function and the convexity depends on whether

$$A^T A - \gamma I \succeq 0$$

can be achieved. Likewise,

$$f \text{ concave} \Longleftrightarrow A^T A - \gamma I \preceq 0$$

Consider SVD $A = U\Sigma V^T$. Then
$$A^T A = V\Sigma^2 V^T$$

This is also symmetric eigendecomposition, with $Q = V$ and $\Lambda = \Sigma^2$. Then we have

$$A^T A - \gamma I = V\Sigma^2 V^T - \gamma I = V(\Sigma^2 - \gamma I)V^T$$

This is actually the eigendecomposition of $A^T A - \gamma I$. Hence

$$A^T A - \gamma I \succeq 0 \iff \sigma_i^2 - \gamma \geq 0,\ i = 1, \dots, n$$
$$A^T A - \gamma I \preceq 0 \iff \sigma_i^2 - \gamma \leq 0,\ i = 1, \dots, n$$

Concluding, f is convex if $\gamma \leq \sigma_{\min}^2$, f is concave if $\gamma \geq \sigma_{\max}^2$, and f is neither convex nor concave, otherwise.

(c) Let $g_y(X) = \frac{y^T X y}{y^T y} = \mathrm{tr}(Xyy^T)/y^T y$. $g_y(X)$ is linear in X. Since $f(X) = \inf_{y\in\mathbf{R}^n, y\neq 0} g_y(X)$ is a point-wise infimum of a concave (linear) function, f is concave.

(d) Let $g(x, \omega) = \log \sum_{i=1}^n x_i cos((i-1)\omega)$. For each ω, $g(x, \omega)$ is a concave function (in x with domain $\mathbf{dom} f$). Since f is an integration of all $g(x, \omega)$, f is concave.

(e) $p(t)$ is an affine function of x. Therefore $\sup_{t\in[a,b]} p(t)$ is convex in x. Similarly, $\inf_{t\in[a,b]} p(t)$ is concave in x. Therefore $f(x)$ is convex in x.

(f) Consider the function $g(t) = e^{-1/t}$ with domain $t < 0$. We have $g'(t) = \frac{1}{t^2}e^{-1/t}$ and $g''(t) = (-\frac{2}{t^3} + \frac{1}{t^4})e^{-1/t} \geq 0$ on $t \geq 0$. Therefore $g(t)$ is convex. Let $\tilde{g}(t)$ be the extended-value extension of $g(t)$. Then $\tilde{g}(t)$ is nondecreasing. By composition rule $g(f_i(x))$ is convex. Therefore $f(x) = \sum_{i=1}^m g(f_i(x))$ is convex.

Q2. Let $p(x)$ and $q(x)$ be functions where $p(x) > 0$ for $x \in S$, $q(x) > 0$ for $x \in S$, $\int_S p(x)dx = 1$ and $\int_S q(x)dx = 1$. Show that

$$\int_S p(x)\log q(x)dx \leq \int_S p(x)\log p(x)dx.$$

Solution:

We need a little variation of Jensen's inequality: For a convex $f : \mathbb{R} \to \mathbb{R}$,

$$f(\sum_{i=1}^n \theta_i g(x_i)) \leq \sum_{i=1}^n \theta_i f(g(x_i))$$

for $\sum_{i=1}^n \theta_i = 1$, and for any function g. Likewise,

$$f(\int_S g(x)p(x)dx) \leq \int_S p(x)f(g(x))dx$$

Now, let $f(x) = -\log x$, and $g(x) = q(x)/p(x)$. By the above Jensen's inequality,

$$-\log\Big(\int_S p(x)\frac{q(x)}{p(x)}dx\Big) \leq -\int_S p(x)\log\frac{q(x)}{p(x)}dx$$
$$\iff -\log 1 \leq \int_S p(x)\log p(x)dx - \int_S p(x)\log q(x)dx$$
$$\iff \int_S p(x)\log q(x)dx \leq \int_S p(x)\log p(x)dx$$

Note: It is interesting to see when the inequality can be achieved. This requires

$$\int_S p(x) \log \frac{q(x)}{p(x)} dx = 0.$$

Since $p(x) > 0,\ \forall x \in S$, the above condition holds iff

$$\log \frac{q(x)}{p(x)} = 0,\ \forall x \in S \Longleftrightarrow p(x) = q(x),\ \forall x \in S$$

Q3. (15%) Show that the following function is convex.

$$f(x) = x^T (A(x))^{-1} x, \qquad \mathbf{dom} f = \{x \mid A(x) \succ 0\},$$

where $A(x) = A_0 + A_1 x_1 + \ldots + A_n x_n \in \mathbf{S}^n$, and $A_i \in \mathbf{S}^n$, $i = 1, \ldots, n$. Hint: You are allowed to use a special form of Schur complement, described as follows: Suppose $A \succ 0$. Then

$$\begin{bmatrix} A & b \\ b^T & c \end{bmatrix} \succeq 0 \Longleftrightarrow c - b^T A^{-1} b \geq 0.$$

Solution: We will show the epigraph $\{(x,t) \mid x^T (A(x))^{-1} x \leq t, A(x) \succ 0\}$ is convex. Using the Schur complement, the epigraph can be written as

$$\left\{ (x,t) \;\middle|\; \begin{bmatrix} A(x) & x \\ x^T & t \end{bmatrix} \succeq 0, A(x) \succ 0 \right\}.$$

$\begin{bmatrix} A(x) & x \\ x^T & t \end{bmatrix}$ and $A(x)$ are affine mapping of (x,t). So the epigraph is convex.

Q4. Let $f_0, \ldots, f_n : \mathbf{R} \to \mathbf{R}$ be given continuous functions. We consider the problem of approximating f_0 as linear combination of $f_1, \ldots, f_n$. For $x \in \mathbf{R}^n$, we say that $f = x_1 f_1 + \ldots + x_n f_n$ approximates f_0 with tolerance $\epsilon > 0$ over the interval $[0,\ T]$ if $|f(t) - f_0(t)| \leq \epsilon$ for $0 \leq t \leq T$. Now we choose a fixed tolerance $\epsilon > 0$ and define the approximation width as the largest T such that f approximates f_0 over the interval $[0, T]$:

$$W(x) = \sup\{T \mid |x_1 f_1(t) + \ldots + x_n f_n(t) - f_0(t)| \leq \epsilon \text{ for } 0 \leq t \leq T\}.$$

Show that W is quasiconcave.

Solution:

We need to show the set $R_\alpha = \{x \mid W(x) \geq \alpha\}$ is convex for any α. We have

$$\begin{aligned} & W(x) \geq \alpha \\ \Longleftrightarrow & \text{for all } T' < \alpha, \text{there exists } T \geq T' \text{ such that } |x_1 f_1(t) + \ldots + x_n f_n(t) - f_0(t)| \leq \epsilon \text{ for } t \in [0, T] \\ \Longleftrightarrow & \text{it holds true that } |x_1 f_1(t) + \ldots + x_n f_n(t) - f_0(t)| \leq \epsilon \text{ for } t \in [0, \alpha) \end{aligned}$$

Therefore we can rewrite R_α as

$$\begin{aligned} R_\alpha &= \{x \mid |x_1 f_1(t) + \ldots + x_n f_n(t) - f_0(t)| \leq \epsilon, t \in [0, \alpha)\} \\ &= \bigcap_{t \in [0,\alpha)} \{x \mid |x_1 f_1(t) + \ldots + x_n f_n(t) - f_0(t)| \leq \epsilon\}. \end{aligned}$$

R_α is the intersection of convex sets and thus is convex.

Q5. Show that $f(X) = (\mathbf{det}X)^{1/n}$ is concave on $\mathbf{S}^n_{++}$. (Hint: Use the same method of proof for showing the concavity of the log determinant function; see page 74 of the textbook)

Solution: Let $U \in \mathbf{S}^n_{++}$ and $V \in \mathbf{S}^n$. We will show that

$$g(t) = (\mathbf{det}U + tV)^{1/n}, \quad \mathbf{dom}\; g = \{t \in \mathbf{R} \mid U + tV \in \mathbf{S}^n_{++}\}$$

is concave. We can write $g(t)$ as

$$\begin{aligned}
g(t) &= (\mathbf{det}(U + tV))^{1/n} \\
&= (\mathbf{det}(U^{1/2}U^{1/2} + tV))^{1/n} && \text{(as } U \text{ is PD, } U^{1/2} \in \mathbf{S}^n_{++} \text{ exists and is invertible)} \\
&= (\mathbf{det}U)^{1/n}(\mathbf{det}(I + tU^{-1/2}VU^{-1/2}))^{1/n} \\
&= (\mathbf{det}U)^{1/n}(\mathbf{det}(I + tQ\Lambda Q^T))^{1/n} && \text{(use EVD } Q\Lambda Q^T = U^{-1/2}VU^{-1/2}\text{)} \\
&= (\mathbf{det}U)^{1/n}\left(\prod_{i=1}^{n}(1 + t\lambda_i)\right)^{1/n}
\end{aligned}$$

Therefore, $g(t)$ is the a (scaled) composition of geometric mean (which is concave) and affine function, and thus concave.