HW 1

Sunday, February 9, 2020 5:41 PM

2.2. Solution: ① If a set S is a convex set,

$\because$ The intersection of two convex sets is convex

$\because$ any line is convex

$\therefore$ The intersection of S and any line is convex.

② Conversely, we first suppose that the intersection of S and any line is convex.

Then we take $\forall x_1, x_2 \in S$, so the intersection of S and line that can through x_1 and x_2 is convex.

$\therefore$ convex combinations of x_1 and x_2 belong to S

$\therefore$ S is a convex set.

2.3. Solution: We can find the midpoint of $a, b \Rightarrow \frac{1}{2}(a+b)$

 then we find another midpoint of a and $(\frac{1}{2}a+\frac{1}{2}b) \Rightarrow \frac{1}{2}a+\frac{1}{2}(\frac{1}{2}a+\frac{1}{2}b)$

after this, we just need to find the midpoint close to a or b.

This midpoint can be expressed as: $\theta_k a + (1-\theta_k) b$

$\theta_k = c_1 2^{-1} + c_2 2^{-2} + \cdots + c_k 2^{-k}$

If we choose the a, b on the border of this set (C is closed).

then we can put k to infinity

$\lim_{k\to\infty} \theta_k = 1$, then $\theta_k \in [0,1]$

hence $\theta x + (1-\theta) y \in C$ for all $\theta \in [0,1]$, $x, y \in C$, then C is convex.

2.10 (a) Solution: As we know, The α-sublevel set of a function $f: R^n \to R$ is defined as $C_\alpha = \{x \in \mathbf{dom} f \mid f(x) \le \alpha\}$

sublevel sets of a convex function are convex for any α.

for this context, the α is 0.

Then we only need to let the function $x^T A x + b^T x + c$ be convex function.

It's a quadratic function, to satisfy it's convex, we need to apply second-order conditions on it. $\nabla^2 f(x) \succeq 0$

so, only if $A \succeq 0$, C is convex.

2.12. (a) A slab is a set of the form $\{x \in R^n \mid \alpha \le a^T x \le \beta\}$.

We can know that it's the intersection of two halfspaces.

halfspace is convex, then slab is convex.

(b). For the form $\{x \in R^n \mid \alpha_i \le x_i \le \beta_i, i = 1, \cdots n\}$

It's the finite intersection of halfspaces, then it's convex.

(c) For the form $\{x \in R^n \mid a_1^T x \le b_1, a_2^T x \le b_2\}$

It's also a intersection of two halfspaces, then it's convex.

(d) For this set $\{x \mid \|x - x_0\|_2 \le \|x-y\|_2$ for all $y \in S\}$, $S \subseteq R^n$

So for a fitted y, $\|x - x_0\|_2 \le \|x-y\|_2$

$\Leftrightarrow (x-x_0)^T (x-x_0) \le (x-y)^T (x-y)$

$\Leftrightarrow x^T x - 2x_0^T x + x_0^T x_0 \le x^T x - 2y^T x + y^T y$

$\Leftrightarrow 2(y-x_0)^T x \le y^T y - x_0^T x_0$

This is a halfspace, then for all $y \in S$, this set is intersection of halfspaces, then it's a convex set.

2.21. Solution: The set $(a, b) \in R^{n+1}$

$a^T x \le b \quad x \in C$

$a^T x \ge b \quad x \in D.$

First, we can find that origin can satisfy both conditions.

Then, from these two conditions we can say that these form an intersection of halfspaces.

Overall, this is a convex cone.

Prove:

① Prove the convexity of e^{ax} on $\mathbb{R}$ for any $a \in R$.

Solution: Consider the second order condition.

$\nabla^2 f(x) = a^2 e^{ax}$ for any $a \in R$, $\nabla^2 f(x) \succeq 0$

then e^{ax} on $\mathbb{R}$ is convex

② Prove the concavity of $\log x$ on R_{++}

Solution: Consider the second order condition.

$\nabla^2 f(x) = \dfrac{-\ln(10)}{x^2 \ln^2(10)} < 0$, then it's concave.

③ Prove the convexity of $-\log\det X$ on S^n_{++}

Solution: Fact: a function is convex if and only if its restriction to any line is convex.

To prove the convexity of $-\log\det X$ on S^n_{++}

we can first apply $\log\det X$

consider the line: $\{x \mid x + tv\}$ $g(t) = f(x+tv) = \log\det(x+tv)$

X is positive definite metrix, then $X^{\frac{1}{2}} X^{\frac{1}{2}} = X$

$$g(t) = \log\det(X^{\frac{1}{2}} X^{\frac{1}{2}} + t X^{\frac{1}{2}} X^{-\frac{1}{2}} V X^{-\frac{1}{2}} X^{\frac{1}{2}})$$

$$= \log\det(X^{\frac{1}{2}} (I + t X^{-\frac{1}{2}} V X^{-\frac{1}{2}}) X^{\frac{1}{2}})$$

$$= \log(\det(X) \det(I + t X^{-\frac{1}{2}} V X^{-\frac{1}{2}}))$$

$$= \log\det(X) + \log\det(I + t X^{-\frac{1}{2}} V X^{-\frac{1}{2}})$$

Assume the eigenvalues of $X^{-\frac{1}{2}} V X^{-\frac{1}{2}}$ are $\lambda_1 \lambda_2 \cdots \lambda_d$, then

$$\log\det(I + t X^{-\frac{1}{2}} V X^{-\frac{1}{2}}) = \log \prod_{i=1}^{d} (1 + t\lambda_i) = \sum_{i=1}^{d} \log(1 + t\lambda_i)$$

then $g(t) = \log\det(X) + \sum_{i=1}^{d} \log(1 + t\lambda_i)$

then $-g(t) = -\log\det(X+tV) = -\log\det(X) - \sum_{i=1}^{d} \log(1 + t\lambda_i)$

Consider the second order condition:

$$-g''(t) = \sum_{i=1}^{d} \frac{\lambda_i^2}{(1+t\lambda_i)^2} \ge 0$$

Thus, $-g(t)$ is convex, so is $-f(x)$ which is $-\log\det X$ on S^n_{++}

3.11. Solution: $f: R^n \to R$ is a differentiable convex function.

then $f(y) \ge f(x) + \nabla f(x)^T (y-x)$, for all $x, y \in \mathbf{dom} f$.

$f(x) \ge f(y) + \nabla f(y)^T (x-y)$.

then $(\nabla f(x) - \nabla f(y))^T (x-y) \ge 0$. satisfy the condition of monotone.

then gradient ∇f is monotone.

The converse is not true.

according to the context, $\psi = \nabla f$

Suppose we have $\psi(x) = \begin{bmatrix} x_1 \\ 2x_1 + x_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$

$\psi(y) = \begin{bmatrix} y_1 \\ 2y_1 + y_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$

ψ is monotone because: $[\psi(x) - \psi(y)]^T (x-y)$

$$= \begin{bmatrix} x_1 - y_1 \\ 2(x_1-y_1) + x_2 - y_2 \end{bmatrix}^T = \left(\begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} x_1 - y_1 \\ x_2 - y_2 \end{bmatrix} \right)^T (x-y)$$

$$= (x-y)^T \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} (x-y)$$

$$= [(x_1 - y_1) + (x_2 - y_2)]^2 \ge 0 \text{ for all } x, y.$$

However, there does not exist a function: $f: R^2 \to R$ that $\psi(x) = \nabla f(x)$

because if $\nabla f(x) = \psi(x)$, then $\dfrac{\partial^2 f}{\partial x_1 \partial x_2} = \dfrac{\partial \psi_1}{\partial x_2} = 0$ $\quad \dfrac{\partial^2 f}{\partial x_1 \partial x_2} = \dfrac{\partial \psi_2}{\partial x_1} = 2.$

It's impossible.

3.18 Solution: (a) define $g(t) = f(Z + tV)$, where $Z \succ 0$, $V \in S^n$

$g(t) = \mathrm{tr}((Z + tV)^{-1})$

$= \mathrm{tr}(Z^{-1}(I + tZ^{-1/2} V Z^{-1/2})^{-1})$

$Z^{-\frac{1}{2}} V Z^{-\frac{1}{2}} = Q \Lambda Q^T$ is the Schur decomposition.

$Q^T Q = Q Q^T = I$ $\quad \Lambda = \mathrm{diag}(\lambda_1, \lambda_2 \cdots \lambda_n)$

$g(t) = \mathrm{tr}(Z^{-1}(I + t Q \Lambda Q^T)^{-1})$

$= \mathrm{tr}(Z^{-1} Q (I + t\Lambda)^{-1} Q^T)$

$= \mathrm{tr}(Q^T Z^{-1} Q (I + t\Lambda)^{-1})$

$= \sum_{i=1}^{n} (Q^T Z^{-1} Q)_{ii} (1 + t\lambda_i)^{-1}$

$g''(t) = \dfrac{\lambda_i (2\lambda_i^2 t + 2\lambda_i)}{(1+t\lambda_i)^{4}} > 0$, then it's convex.

(b). Define: $g(t) = f(Z + tV)$ where $Z \succ 0$, $V \in S^n$

$g(t) = (\det(Z + tV))^{\frac{1}{n}}$

$= (\det Z^{\frac{1}{2}} \det(I + t Z^{-\frac{1}{2}} V Z^{-\frac{1}{2}}) \det Z^{\frac{1}{2}})^{\frac{1}{n}}$

$= (\det Z)^{\frac{1}{n}} (\det(I + t Z^{-\frac{1}{2}} V Z^{-\frac{1}{2}}))^{\frac{1}{n}}$

$= (\det Z)^{\frac{1}{n}} (\prod_{i=1}^{n} (1 + t\lambda_i))^{\frac{1}{n}}$

for $\prod_{i=1}^{n} (1 + t\lambda_i)^{\frac{1}{n}}$ is a kind of geometric mean.

then $\nabla^2 f(x) \le 0$, $Z \succ 0$

hence $\nabla^2 g(t) \le 0$, it's concave.