Convex Functions (cont'd)

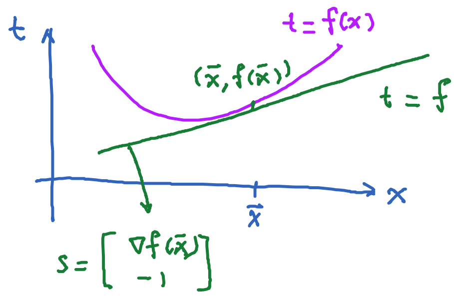
2022年3月3日 15:26

Differentiable Convex functions

Theorem: Let $f : \mathbb{R}^n \to \mathbb{R}$ be differentiable. Then, f is convex iff

$$f(x) \geq \underbrace{f(\bar{x}) + \nabla f(\bar{x})^T (x - \bar{x})}_{\text{affine function}} \quad \forall x, \bar{x} \in \mathbb{R}^n.$$

$x \mapsto c^T x + d$



$$t = f(\bar{x}) + \nabla f(\bar{x})^T (x - \bar{x})$$

$$\Updownarrow$$

$$\underbrace{\begin{bmatrix} \nabla f(\bar{x}) \\ -1 \end{bmatrix}}_{s}^T \begin{bmatrix} x \\ t \end{bmatrix} + f(\bar{x}) - \nabla f(\bar{x})^T \bar{x} = 0$$

$z : s^T z = c$ ← hyperplane

Theorem: Let $f : S \to \mathbb{R}$ be a twice continuously differentiable function on the open convex set $S \subseteq \mathbb{R}^n$. Then,

$$f \text{ is convex on } S \iff \nabla^2 f(x) \succeq 0 \quad \forall x \in S$$

$f(\alpha x + (1-\alpha) y) \leq \alpha f(x) + (1-\alpha) f(y)$

$\forall x, y \in S;\ \alpha \in [0,1]$

Example: Let $f(x) = \frac{1}{2} x^T A x + c^T x$ with $A \in \mathcal{S}^n$ and $c \in \mathbb{R}^n$.

($f : \mathbb{R}^n \to \mathbb{R}$; $\mathbb{R}^n$ is open and convex)

$$\nabla f(x) = Ax + c, \quad \nabla^2 f(x) = A$$

Hence, f is convex $\iff A \succeq 0$

Example: Let $f(x, y) = x^2 - y^2$, $S = \mathbb{R} \times \{0\} = \{(x, 0) : x \in \mathbb{R}\}$

$f : \mathbb{R}^2 \to \mathbb{R}$

① f is convex on S (Verify)

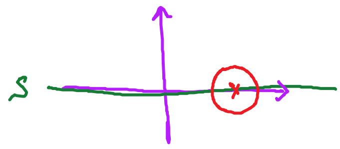
② $\nabla f(x, y) = 2 \begin{bmatrix} x \\ -y \end{bmatrix}$; $\nabla^2 f(x, y) = \begin{bmatrix} 2 & 0 \\ 0 & -2 \end{bmatrix} \not\succeq 0$

Note that S is convex but not open.

Recall: $x \in \text{int}(S)$

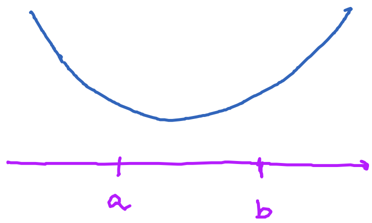
if $\exists \varepsilon > 0$ s.t.

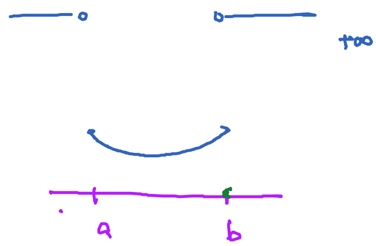
$B(x, \varepsilon) \subseteq S$



S is open iff $S = \text{int}(S)$

$$\min \; f(x) \quad \text{s.t.} \; x \in P \qquad\qquad \min \; f(x) + \mathbb{1}_P(x)$$





Example: Let $f : \underset{\text{pd, open set}}{S^n_{++}} \to \mathbb{R}$ be given by $f(X) = -\ln \det(X)$

(Recall: $\det(X) = \prod_{i=1}^{n} \lambda_i(X)$) Is f convex?

Approach 1: By calculus,

$$\nabla f(X) = -X^{-1}, \qquad \nabla^2 f(X) = X^{-1} \otimes X^{-1}$$

Kronecker product

Recall: $\underset{m\times n}{A} \otimes \underset{p\times q}{B} = \begin{bmatrix} a_{11}B & \cdots & a_{1n}B \\ \vdots & & \\ a_{m1}B & \cdots & a_{mn}B \end{bmatrix} \in \mathbb{R}^{mp \times nq}$

Fact: If $A \succ 0$, then $A \otimes A \succ 0$.

Hence, if $X \succ 0$, then $X^{-1} \succ 0$ and $X^{-1} \otimes X^{-1} \succ 0$

Approach 2: By restriction on lines

Let $X_0 \in S^n_{++}$ and $H \in S^n$ be given. Define

$$t \mapsto \tilde{f}_{X_0, H}(t) = -\ln \det(X_0 + tH)$$

Want: $\tilde{f}_{X_0, H}$ is convex for all $X_0 \in S^n_{++}$, $H \in S^n$.

Since $X_0 \in S^n_{++}$, we can write $X_0 = \underset{\text{eigendecomposition}}{\uparrow}\ U \Lambda U^T = \underbrace{(U \Lambda^{1/2} U^T)}_{X_0^{1/2}} \underbrace{(U \Lambda^{1/2} U^T)}_{X_0^{1/2}}$.

orthogonal (U), diagonal (Λ)

Then,

$$\tilde{f}_{X_0,H} = -\ln \det \left[X_0^{1/2} \left(I + t \underset{\substack{\shortparallel \\ (X_0^{1/2})^{-1}}}{X_0^{-1/2}} H X_0^{-1/2} \right) X_0^{1/2} \right]$$

$$= -2 \ln \det (X_0^{1/2}) - \ln \det \left(I + t \underbrace{X_0^{-1/2} H X_0^{-1/2}}_{\in S^n} \right)$$

Let $\lambda_1, \lambda_2, \dots, \lambda_n$ be the eigenvalues of $X_0^{-1/2} H X_0^{-1/2}$. Then, the eigenvalues of $I + t X_0^{-1/2} H X_0^{-1/2}$ are $\{1 + t\lambda_i\}_{i=1}^n$.

Hence,

$$-\ln \det (I + t X_0^{-1/2} H X_0^{-1/2}) = \sum_{i=1}^n - \ln (1 + t\lambda_i)$$

It remains to show that $t \mapsto -\ln(1 + t\underline{\underline{\lambda}})$ is convex. (Exercise)