Lecture #3

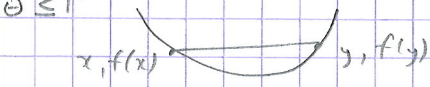
19 Oct 2016

§3 Convex Functions | Lecture Slides

Contents:
- basic prop & e.g.
- operations that preserve convexity
- the conjugate f^n
- quasiconvex f^ns
- log-concave & log-convex f^ns
- convexity wrt generalized inequalities

Def^{n}s

Def^n: Convex := $f: \mathbb{R}^n \to \mathbb{R}$ is convex iff dom f is convex & $f(\theta x + (1-\theta)y) \leq \theta f(x) + (1-\theta) f(y)$ $\forall\, x, y \in \text{dom} f$, $0 \leq \theta \leq 1$



Def^n: Concave := if $-f$ is convex.

Def^n: Strictly convex := if dom f is convex & $f(\theta x + (1-\theta)y) < \theta f(x) + (1-\theta) f(y)$ $\forall\, x, y \in \text{dom} f$, $x \neq y$, $0 < \theta < 1$

Examples on $\mathbb{R}$

e.g. convex:
- affine: $ax+b$ on $\mathbb{R}$, $a, b \in \mathbb{R}$
- exponential: e^{ax}, for any $a \in \mathbb{R}$
- powers: x^α on $\mathbb{R}_{++}$, $\alpha \geq 1$ or $\alpha \leq 0$
- powers of abs: $|x|^p$ on $\mathbb{R}$, $p \geq 1$
- neg. entropy: $x \log x$ on $\mathbb{R}_{++}$

concave:
- affine: $ax+b$ on $\mathbb{R}$, $a, b \in \mathbb{R}$
- powers: x^α on $\mathbb{R}_{++}$ $0 \leq \alpha \leq 1$
- log: $\log x$ on $\mathbb{R}_{++}$

Examples on $\mathbb{R}^n$ & $\mathbb{R}^{m\times n}$

affine f^n are convex & concave; all norms are convex

e.g. on $\mathbb{R}^n$: ~~affine f^ns~~: $f(x) = a^T x + b$

→ norms: $\|x\|_p = \left(\sum_i |x_i|^p\right)^{1/p}$ for $p \geq 1$

$\|x\|_\infty = \max_k |x_k|$

e.g. on $\mathbb{R}^{m\times n}$

→ affine f^n $f(X) = \text{tr}(A^T X) + b = \sum_{i=1}^m \sum_{j=1}^n A_{ij} X_{ij} + b$

→ spectral (max. singular value) norm $f(X) = \|X\|_2 = \sigma_{max}(X) = (\lambda_{max} X^T X)^{1/2}$

Restriction of a convex f^n to a line

claim:

$f: \mathbb{R}^n \to \mathbb{R}$ is convex iff the f^n $g: \mathbb{R} \to \mathbb{R}$

$g(t) = f(x + tv)$ $\quad$ dom$(g) = \{t \mid x + tv \in \text{dom} f\}$

is convex (in t) for any $x \in$ dom(f), $v \in \mathbb{R}^n$

Remark: Can check convexity of f by checking convexity of one variable f^n functions

e.g. $f: S^n \to \mathbb{R}$ with $f(X) = \log \det X$, dom $X = S^n_{++}$

$g(t) = \log \det(X + tV) = \log \det X + \log \det(\mathbb{1} + tX^{-1/2} V X^{-1/2})$

$= \log \det X + \sum \log(1 + t\lambda_i)$

with λ_i as eigenvalues of $X^{-1/2} V X^{-1/2}$

g is concave in t ($\forall\, X \succ 0$, V); hence f is concave

Extended-value extensions

Def^n: $\tilde{f}$:=

$\tilde{f} = f(x)$, $x \in$ dom f

$\tilde{f} = \infty$ $\quad x \notin$ dom f

often helps.

e.g. $0 \leq \theta \leq 1 \Rightarrow \tilde{f}(\theta x + (1-\theta)y) \leq \theta \tilde{f}(x) + (1-\theta)\tilde{f}(y)$

(as an inequality in $\mathbb{R} \cup \{\infty\}$) means the same as

(1) domain f is convex & (2) $\forall\, x, y \in$ dom f

$0 \leq \theta \leq 1 \Rightarrow$

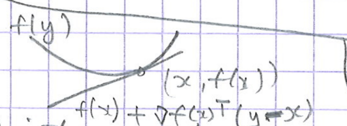
$f(\theta x + (1-\theta)y) \leq \theta f(x) + (1-\theta) f(y)$

First-order condition

Def^n: differentiable := if dom f is open & the gradiant

$$\nabla f(x) = \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \cdots \frac{\partial f}{\partial x_n}\right)$$

exists at each $x \in$ dom f



Claim:

(1^{st} order condition:) a differentiable f with convex domain is convex iff

$f(y) \geq f(x) + \nabla f(x)^T (y - x)$ $\forall\, x, y \in$ dom f

NB: First-order approximation of f is a global underestimation.

Second-order conditions

Def^n: twice differentiable := if dom f is open & the hessian $\nabla^2 f(x) \in S^n$ where $\nabla^2 f(x)_{ij} = \dfrac{\partial^2 f}{\partial x_i \partial x_j}$ exists at each $x \in$ dom f.

Claim:

(2^{nd} order condition:) a twice differentiable f with convex domain

1) is convex iff $\nabla^2 f(x) \succeq 0$ $\forall\, x \in$ dom f

2) if $\nabla^2 f(x) \succ 0$ $\forall\, x \in$ dom f, $\Rightarrow f$ is strictly convex.

Examples

(1) Quadratic f^n:

$f(x) = \frac{1}{2} x^T P x + q^T x + r$

(with $P \in S^n$)

$\nabla f(x) = Px + q$ $\quad \nabla^2 f(x) = P$

convex if $P \succeq 0$

(2) least squares objective :=

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

$\nabla f(x) = 2A^T(Ax - b)$

$\nabla^2 f(x) = 2A^T A$

convex for any A.

(3) Quadratic-over-linear:

$f(x, y) = x^2 / y$

$\nabla^2 f(x,y) = \frac{2}{y^3} \begin{bmatrix} y \\ -x \end{bmatrix} \begin{bmatrix} y \\ -x \end{bmatrix}^T \succeq 0$

convex $\forall\, y > 0$ $\Leftarrow$

(4) log sum exp:

$f(x) = \log \sum_k e^{x_k}$

$\nabla^2 f(x) = \frac{1}{\mathbb{1}^T z} \text{diag}(z) - \frac{z z^T}{(\mathbb{1}^T z)^2}$ $\quad (z_k = e^{x_k})$

claim: $\nabla^2 f(x) \succeq 0$

"proof": $v^T \nabla^2 f(x)\, v = \dfrac{(\sum z_k v_k^2)(\sum_k z_k) - (\sum_k v_k z_k)^2}{(\sum_k z_k)^2} \geq 0$

since $(\sum_k v_k z_k)^2 \leq (\sum_k z_k v_k^2)(\sum z_k)$ (from Cauchy-Schwarz inequality)

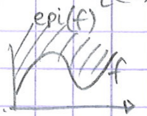
(5) Geometric Mean: $f(x) = (\prod_k x_k)^{1/n}$ on $\mathbb{R}^n_{++}$ is concave (proof is similar to log-sum)

Epigraph and sublevel set

Def^n: α-sublevel set := of $f: \mathbb{R}^n \to \mathbb{R}$ is
$C_\alpha := \{x \in \text{dom} f \mid f(x) \leq \alpha\}$

claim: sublevel sets of convex f^{n}s are convex (converse is false)

Def^n: Epigraph of $f: \mathbb{R}^n \to \mathbb{R}$:= epi(f) := $\{(x,t) \in \mathbb{R}^{n+1} \mid x \in \text{dom} f,\ f(x) \leq t\}$



claim: f is convex iff ($\Leftrightarrow$) epi(f) is a convex set.

Jensen's Inequality

Recall: if f is convex, for $0 \leq \theta \leq 1$,
$f(\theta x + (1-\theta) y) \leq \theta f(x) + (1-\theta) f(y)$

Jensen's: if f is convex, then

$$f(Ez) \leq E f(z)$$

for any random variable z.

NB: you recover the original with $Prob(z = x) = \theta$ & $Prob(z = y) = 1-\theta$

Operations that preserve convexity

Aim: Practical methods for establishing convexity.

1. Verify def^n
2. for twice diff, $\nabla^2 f(x) \succeq 0$
3. Show by actions that preserve convexity
   (a) non-neg. weighted sum
   (b) composition with affine
   (c) pointwise max & supremum.
   (d) composition.
   (e) minimization
   (f) perspective

Positive weighted & composition with affine f^n (obvious ones)

(a1) non-neg. multiple: αf is convex if f is convex
(a2) sum: $f_1 + f_2$ (extends to integrals)
(a3) composition with affine: $f(Ax+b)$

e.g. log barrier for linear ineq: $f(x) = -\sum \log(b_i - a_i^T x)$

e.g. norm of affine f^n: $f(x) = \|Ax+b\|$

Pointwise Max

claim: $f_1(x) \ldots f_m(x)$ are convex $\Rightarrow$ $\max(f_1(x) \ldots f_m(x))$ is convex

e.g. sum of r largest components of $x \in \mathbb{R}^n$
$f(x) = x_{[1]} + x_{[2]} + \cdots + x_{[r]}$ is convex.
where $x_{[i]}$ is the ith largest component of x

Pointwise Supremum

claim: If $f(x,y)$ is convex in x for each $y \in \mathcal{A}$, then
$g(x) = \sup_{y \in \mathcal{A}} f(x,y)$
is convex

e.g. (1) support f^n of set C: $S_C(x) = \sup_{y \in C} y^T x$ is convex.

(2) distance to farthest point in a set C
$f(x) = \sup_{y \in C} \|x - y\|$

(3) max. eigenvalue of a symmetric matrix; for $X \in S^n$
$\lambda_{max}(X) = \sup_{\|y\|_2 = 1} y^T X y$

Vector Composition

claim: $g: \mathbb{R}^n \to \mathbb{R}^k$
$h: \mathbb{R}^k \to \mathbb{R}$

$f(x) = h(g(x))$
$= h(g_1(x), g_2(x), \ldots g_k(x))$

is convex, if
(1) g_i convex, h convex, $\tilde{h}$ nondecr in each arg
or
(2) g_i concave, h convex, $\tilde{h}$ nonincr in each arg

"proof": $f''(x) = g'(x)^T \nabla^2 h(g(x))\, g'(x) + \nabla h(g(x))^T g''(x)$

Composition with scalar f^{n}s

claim: $f(x) = h(g(x))$ is convex if
(a) g convex, h convex, $\tilde{h}$ nondecreasing
(b) g concave, h convex, $\tilde{h}$ nonincr.

"proof": $f''(x) = h''(g(x))\, g'(x)^2 + h'(g(x)) \cdot g''(x)$

NB: Monotonicity must hold for extended-valued $\tilde{h}$.

e.g. $\exp(g(x))$ is convex if g is convex
$1/g(x)$ is convex if g is concave & positive

e.g. (1) $\sum_i \log g_i(x)$ is concave if g_i is concave & positive

(2) $\log \sum_i \exp(g_i(x))$ is convex if $g(i)$ are convex