3. Convex functions

- basic properties and examples
- operations that preserve convexity
- the conjugate function
- quasiconvex functions
- log-concave and log-convex functions
- convexity with respect to generalized inequalities

Definition

$f : \mathbf{R}^n \to \mathbf{R}$ is convex if $\mathbf{dom}\, f$ is a convex set and

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

for all $x, y \in \mathbf{dom}\, f$, $0 \le \theta \le 1$

- f is concave if $-f$ is convex
- f is strictly convex if $\mathbf{dom}\, f$ is convex and

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

  for $x, y \in \mathbf{dom}\, f$, $x \ne y$, $0 < \theta < 1$

Examples on R

convex:

- affine: $ax + b$ on $\mathbf{R}$, for any $a, b \in \mathbf{R}$
- exponential: e^{ax}, for any $a \in \mathbf{R}$
- powers: x^α on $\mathbf{R}_{++}$, for $\alpha \ge 1$ or $\alpha \le 0$
- powers of absolute value: $|x|^p$ on $\mathbf{R}$, for $p \ge 1$
- negative entropy: $x \log x$ on $\mathbf{R}_{++}$

concave:

- affine: $ax + b$ on $\mathbf{R}$, for any $a, b \in \mathbf{R}$
- powers: x^α on $\mathbf{R}_{++}$, for $0 \le \alpha \le 1$
- logarithm: $\log x$ on $\mathbf{R}_{++}$

Examples on $\mathbf{R}^n$ and $\mathbf{R}^{m \times n}$

affine functions are convex and concave; all norms are convex

examples on $\mathbf{R}^n$

- affine function $f(x) = a^T x + b$
- norms: $\|x\|_p = (\sum_{i=1}^n |x_i|^p)^{1/p}$ for $p \ge 1$; $\|x\|_\infty = \max_k |x_k|$

examples on $\mathbf{R}^{m \times n}$ ($m \times n$ matrices)

- affine function

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

- spectral (maximum singular value) norm

$$f(X) = \|X\|_2 = \sigma_{\max}(X) = (\lambda_{\max}(X^T X))^{1/2}$$

Restriction of a convex function to a line

$f : \mathbf{R}^n \to \mathbf{R}$ is convex if and only if the function $g : \mathbf{R} \to \mathbf{R}$,

$$g(t) = f(x + tv), \qquad \mathbf{dom}\, g = \{t \mid x + tv \in \mathbf{dom}\, f\}$$

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

can check convexity of f by checking convexity of functions of one variable

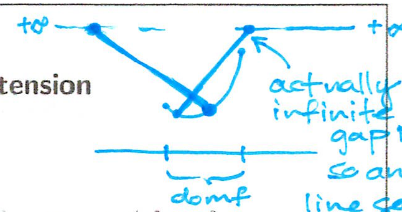
example. $f : \mathbf{S}^n \to \mathbf{R}$ with $f(X) = \log\det X$, $\mathbf{dom}\, f = \mathbf{S}^n_{++}$

$$\begin{aligned} g(t) = \log\det(X + tV) &= \log\det X + \log\det(I + tX^{-1/2}VX^{-1/2}) \\ &= \log\det X + \sum_{i=1}^{n} \log(1 + t\lambda_i) \end{aligned}$$

where λ_i are the eigenvalues of $X^{-1/2}VX^{-1/2}$

g is concave in t (for any choice of $X \succ 0$, V); hence f is concave

Extended-value extension



actually infinite gap here, so any line seg connecting 2 pts is convex in domf

extended-value extension $\tilde{f}$ of f is

$$\tilde{f}(x) = f(x), \quad x \in \mathbf{dom}\, f, \qquad \tilde{f}(x) = \infty, \quad x \notin \mathbf{dom}\, f$$

• $f(x)$ = undef for $x \notin$ domf

often simplifies notation; for example, the condition

$$0 \le \theta \le 1 \quad \Longrightarrow \quad \tilde{f}(\theta x + (1-\theta)y) \le \theta \tilde{f}(x) + (1-\theta)\tilde{f}(y)$$

(as an inequality in $\mathbf{R} \cup \{\infty\}$), means the same as the two conditions

$x, y \in \mathbf{R} \cup \{\infty\}$

- $\mathbf{dom}\, f$ is convex
- for $x, y \in \mathbf{dom}\, f$,

$$0 \le \theta \le 1 \quad \Longrightarrow \quad f(\theta x + (1-\theta)y) \le \theta f(x) + (1-\theta)f(y)$$

• For concave, $\tilde{f}(x) = -\infty$, $x \notin$ domf

First-order condition

f is **differentiable** if $\mathbf{dom}\, f$ is open and the gradient

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

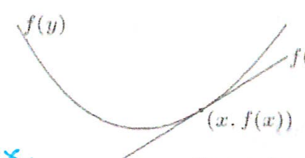
exists at each $x \in \mathbf{dom}\, f$

not rigorous

1st-order condition: differentiable f with convex domain is convex iff

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

• func varies wrt y
• RHS is linear approx on $\mathbb{R}^n$



• x is the point of eval. of 1st order Taylor poly.

first-order approximation of f is global underestimator

If f convex,
• From local info ($\nabla f(x)$), we can guarantee global bound (bc 1st order approx always underestimates)

Second-order conditions

f is **twice differentiable** if $\mathbf{dom}\, f$ is open and the Hessian $\nabla^2 f(x) \in \mathbf{S}^n$,

$$\nabla^2 f(x)_{ij} = \frac{\partial^2 f(x)}{\partial x_i \partial x_j}, \quad i, j = 1, \ldots, n,$$

exists at each $x \in \mathbf{dom}\, f$

2nd-order conditions: for twice differentiable f with convex domain

- f is convex if and only if • 2 way

$$\nabla^2 f(x) \succeq 0 \quad \text{for all } x \in \mathbf{dom}\, f$$

• Hessian of f(x) is PSD for all $x \in$ domf

- if $\nabla^2 f(x) \succ 0$ for all $x \in \mathbf{dom}\, f$, then f is strictly convex

• NOT 2-way for strict case
• $\nabla^2 x^4$ at $x=0$ equals 0 even though x^4 strictly convex

Operations that preserve convexity

practical methods for establishing convexity of a function (i.e., deconstruct a problem (func.) to figure out whether it's convex)

1. verify definition (often simplified by restricting to a line)
2. for twice differentiable functions, show $\nabla^2 f(x) \succeq 0$
3. show that f is obtained from simple convex functions by operations that preserve convexity
   - nonnegative weighted sum
   - composition with affine function
   - pointwise maximum and supremum
   - composition
   - minimization
   - perspective

- 3 is usually preferred to 1 or 2

- Want to decompose problem into simple convex func and convexity-preserving operations (to determine that it's convex)

Positive weighted sum & composition with affine function

nonnegative multiple: αf is convex if f is convex, $\alpha \geq 0$

sum: $f_1 + f_2$ convex if f_1, f_2 convex (extends to infinite sums, integrals)

composition with affine function: $f(Ax + b)$ is convex if f is convex

examples

- log barrier for linear inequalities (i.e., polyhedron)

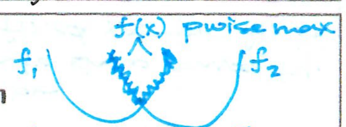
$$f(x) = -\sum_{i=1}^{m} \log(b_i - a_i^T x), \qquad \mathbf{dom}\, f = \{x \mid a_i^T x < b_i, i = 1, \ldots, m\}$$

- $Ax \preceq b$ = intersection of halfspaces $\to b_i - a_i^T x > 0$ = interior of a polyhedron

- (any) norm of affine function: $f(x) = \|Ax + b\|$ is convex

- Appendix has rules for calc. Hessians

Pointwise maximum



if $f_1, \ldots, f_m$ are convex, then $f(x) = \max\{f_1(x), \ldots, f_m(x)\}$ is convex

epi $(\max\{f_1, \ldots, f_m\}) = \bigcap_i \text{epi}\, f_i$

examples

- piecewise-linear function: $f(x) = \max_{i=1,\ldots,m}(a_i^T x + b_i)$ is convex
- sum of r largest components of $x \in \mathbf{R}^n$:

$$f(x) = x_{[1]} + x_{[2]} + \cdots + x_{[r]}$$

  is convex ($x_{[i]}$ is ith largest component of x)

  proof:

$$f(x) = \max\{x_{i_1} + x_{i_2} + \cdots + x_{i_r} \mid 1 \leq i_1 < i_2 < \cdots < i_r \leq n\}$$

  - max of all possible sums of r diff. comp. of x. (i.e., max of piece-wise linear func.)

Pointwise supremum

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

examples

- support function of a set C: $S_C(x) = \sup_{y \in C} y^T x$ is convex
- distance to farthest point in a set C:

$$f(x) = \sup_{y \in C} \|x - y\|$$

- maximum eigenvalue of symmetric matrix: for $X \in \mathbf{S}^n$,

$$\underbrace{\lambda_{\max}(X)}_{\text{linear in } X \text{, so, convex}} = \sup_{\|y\|_2 = 1} \underbrace{y^T X y}_{\text{linear in } X}$$

  → point-wise sup. of inf. linear eq. (1 for each y on unit sphere) quad form + affine

Examples

quadratic function: $f(x) = (1/2)x^TPx + q^Tx + r$ (with $P \in \mathbf{S}^n$)

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

$\nabla x^TPx = 2Px \quad \nabla^2 x^TPx = \nabla 2Px = 2P$

$\nabla q^Tx = q \quad \nabla^2 q^T x = \nabla q = 0$

convex if $P \succeq 0$

least-squares objective: $f(x) = \|Ax - b\|_2^2 = (Ax-b)^T(Ax-b) = (Ax)^TAx - 2b^TAx + b^Tb$

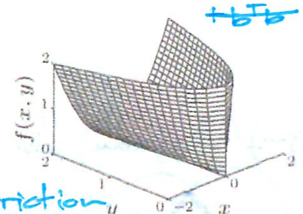
• ex of quad func.

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

convex (for any A) $\quad \nabla x^TA^TAx = 2A^TAx$ $\quad = \cancel{(Ax)^T - 2b^T)Ax + b^Tb}$

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 for $y > 0$

• A convex func. of many variables must be convex over each indiv. variable.
• converse is false.
• If $f(x,y)$ convex in x + y, diag of Hessian are non-neg.

• Can apply restriction of a line theorem to show convexity.
• x^2/y convex for fixed x + for fixed y

log-sum-exp: $f(x) = \log \sum_{k=1}^n \exp x_k$ is convex

• If $\max x_k \gg$ all other $x_k \rightarrow f(x) \approx \max x_k$

$$\nabla^2 f(x) = \underbrace{\frac{1}{\mathbf{1}^Tz}\mathbf{diag}(z)}_{PD} - \frac{1}{(\mathbf{1}^Tz)^2}zz^T \qquad (z_k = \exp x_k) > 0$$

to show $\nabla^2 f(x) \succeq 0$, we must verify that $v^T\nabla^2 f(x)v \geq 0$ for all v:

• i.e., to show a mat. M is PSD, show the quad form is nonneg.

$$v^T\nabla^2 f(x)v = \frac{(\sum_k z_kv_k^2)(\sum_k z_k) - (\sum_k v_kz_k)^2}{(\sum_k z_k)^2} \geq 0$$

just a dot product

since $(\sum_k v_kz_k)^2 \leq (\sum_k z_kv_k^2)(\sum_k z_k)$ (from $\boxed{\text{Cauchy-Schwarz inequality}}$)

Cauchy-Schwarz: $(a^Tb)^2 \leq (a^Ta)(b^Ta)$

$a = v\sqrt{z}$ and $b = \sqrt{z}$

geometric mean: $f(x) = (\prod_{k=1}^n x_k)^{1/n}$ on $\mathbf{R}^n_{++}$ is concave

(similar proof as for log-sum-exp)

$= \{x \in \mathbb{R}^n \mid x_i > 0 \ \forall i\}$

• $\mathbb{R}^n_+$ has ≥ 0

Epigraph and sublevel set

α**-sublevel set** of $f : \mathbf{R}^n \to \mathbf{R}$:

• In estimation, C_α is set of points that are α-plausible.

$$C_\alpha = \{x \in \mathbf{dom}\, f \mid f(x) \leq \alpha\}$$

• sublevel sets of convex functions are convex (converse is false)
• Func. of

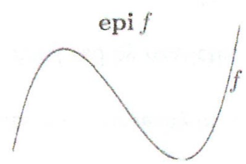
~~α volume~~ of

$C_{\alpha_1} = C_{\alpha_2}$

C_{α_3} the bigger region in xy plane

epigraph of $f : \mathbf{R}^n \to \mathbf{R}$:

$$\mathbf{epi}\, f = \{(x, t) \in \mathbf{R}^{n+1} \mid x \in \mathbf{dom}\, f,\ f(x) \leq t\}$$



f is convex if and only if $\mathbf{epi}\, f$ is a convex set

• Relationship b/t convex func. + convex set
• Very imp. to think about func. + its epigraph together

Jensen's inequality

Let this be a loss func.

• Noise or normal variation ~~Adding noise to~~ on convex func always hurts b/c the 50% w/ steeper slope overtakes other 50%

basic inequality: if f is convex, then for $0 \leq \theta \leq 1$,

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

• Convex combination extends to any # of points
• Equality if all x_i's equal or $f(x)$ is affine

extension: if f is convex, then

• Flip for concave.

$$f(\mathbf{E}\, z) \leq \mathbf{E}\, f(z)$$

• "Full strength"

for any random variable z that is in dom f a.s.

basic inequality is special case with discrete distribution

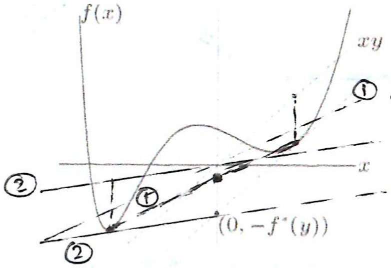
$$\mathbf{prob}(z = x) = \theta, \qquad \mathbf{prob}(z = y) = 1 - \theta$$

"Convex func of an integral/mean/weighted avg $\leq$ i/m/w.a. of convex func."

The conjugate function

the **conjugate** of a function f is

$$f^*(y) = \sup_{x \in \mathbf{dom}\, f} (y^T x - f(x))$$



xy - has slope y

① • As y↓, dashed lines ↓ steep + $-f^*(y)$↑ until ①. This is where $-f^*(y)$ ~~for~~ is max (or $f^*(y)$ is min). ~~The new~~ As y↓ further, the new pt of tangency is ② + $-f^*(y)$↓.

• roughly, $(f^*)^* = f$ if f convex

- f^* is convex (even if f is not)
  - b/c it's a sup of affine (in y) func.
- will be useful in chapter 5 (duality)
- $(f^*)^* = f^{env}$ (convex envelope)
- $\mathrm{epi}(f^{env}) = \mathrm{conv}(\mathrm{epi}\, f)$

examples

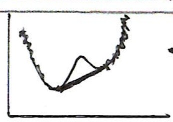
$\sup_{x>0} (xy + \log x)$. $y \geq 0 \to \infty$. Since xy is linear func. of x, if $y<0$, ∃ a maximum wrt x.
$\to y + \frac{1}{x} = 0 \to x = -\frac{1}{y} \to y<0, f^*(y) = -1 - \log(-y)$

- negative logarithm $f(x) = -\log x$

$$\begin{aligned} f^*(y) &= \sup_{x>0}(xy + \log x) \\ &= \begin{cases} -1 - \log(-y) & y < 0 \\ \infty & \text{otherwise} \end{cases} \end{aligned}$$

- strictly convex quadratic $f(x) = (1/2)x^T Q x$ with $Q \in \mathbf{S}^n_{++}$

$$\begin{aligned} f^*(y) &= \sup_x (y^T x - (1/2)x^T Q x) \\ &= \frac{1}{2} y^T Q^{-1} y \end{aligned}$$



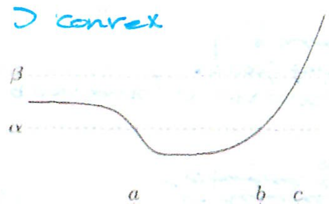
Quasiconvex functions

$f : \mathbf{R}^n \to \mathbf{R}$ is quasiconvex if $\mathbf{dom}\, f$ is convex and the sublevel sets

$$S_\alpha = \{x \in \mathbf{dom}\, f \mid f(x) \leq \alpha\}$$

are convex for all α

- quasiconvex ⊃ convex



- f is quasiconcave if $-f$ is quasiconvex (or equiv., the superlevel sets of f are all convex)
- f is quasilinear if it is quasiconvex and quasiconcave
- In general quasiconvex + convex don't share same preserving or non-preserving operations
  - e.g., quasiconvex closed under monotone transformations but not under sum

Examples

- $\sqrt{|x|}$ is quasiconvex on $\mathbf{R}$ — sublvl set is just an interval
- $\mathrm{ceil}(x) = \inf\{z \in \mathbf{Z} \mid z \geq x\}$ is quasilinear
  - only int-valued func. that is convex is ~~cor~~ a constant
- $\log x$ is quasilinear on $\mathbf{R}_{++}$
- $f(x_1, x_2) = x_1 x_2$ is quasiconcave on $\mathbf{R}^2_{++}$ — associated quad form mat. is $\begin{bmatrix} 0 & 1/2 \\ 1/2 & 0 \end{bmatrix}$ which is indefinite (neg. det.)
- linear-fractional function

$$f(x) = \frac{a^T x + b}{c^T x + d}, \qquad \mathbf{dom}\, f = \{x \mid c^T x + d > 0\}$$

is quasilinear

- distance ratio

$$f(x) = \frac{\|x - a\|_2}{\|x - b\|_2}, \qquad \mathbf{dom}\, f = \{x \mid \|x - a\|_2 \leq \|x - b\|_2\}$$

closer one
farther one

is quasiconvex

Composition with scalar functions

composition of $g : \mathbf{R}^n \to \mathbf{R}$ and $h : \mathbf{R} \to \mathbf{R}$:

$h(g(x)) = h(z)$ with $z = g(x)$ → view monotonicity of h separate from $g(x)$.

$$f(x) = h(g(x))$$

$\tilde{}$ denotes extended val.

(i.e. $f''(x) \geq 0$)

f is convex if
- g convex, h convex, $\tilde{h}$ nondecreasing
- g concave, h convex, $\tilde{h}$ nonincreasing

$\tilde{h}$ nondecreasing (from diff. chain rule below) (twice)

f is concave if g concave, h concave,

- proof (for $n = 1$, differentiable g, h) • proof for only diff func., but h, g need not be diff.

$$f''(x) = h''(g(x))g'(x)^2 + h'(g(x))g''(x)$$

- note: monotonicity must hold for extended-value extension $\tilde{h}$
  - Very imp. For ex, consider $\|Ax-b\|^{1.62}$. Let $h(z) = z^{1.62}$ for $z \geq 0$. Can't use composition thrm. bc $\tilde{h}(z)$ is $+\infty$ for $z<0$, and $\tilde{h}(z)$ NOT nondec. Must define $h(z) = \begin{cases} z^{1.62} & z \geq 0 \\ 0 & \text{otherwise} \end{cases}$ to use thrm.

examples

- $\exp g(x)$ is convex if g is convex
- $1/g(x)$ is convex if g is concave and positive

Vector composition

• Can derive earlier rules by using this
1. Sum of convex func. is convex: $h(g(x)) = \mathbf{1}^T g(x)$ and $g(x)$ = vec of convex func.
2. $\max_i g_i(x)$

composition of $g : \mathbf{R}^n \to \mathbf{R}^k$ and $h : \mathbf{R}^k \to \mathbf{R}$:

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

f is convex if
- g_i convex, h convex, $\tilde{h}$ nondecreasing in each argument
- g_i concave, h convex, $\tilde{h}$ nonincreasing in each argument

• ↑ any component → ↑ max; never change only one arg. at a time

proof (for $n = 1$, differentiable g, h)

$$f''(x) = g'(x)^T \nabla^2 h(g(x)) g'(x) + \nabla h(g(x))^T g''(x)$$

quad form, so sign depends only on convexity/concavity of h

• $g''(x) \succeq 0$ if convex, $g''(x) \preceq 0$ if concave

• ~~$h(z) \succeq 0 \to \nabla h(z)$~~

• $\tilde{h}(z)$ ↗ in each arg → nondec → $\nabla h(z) \succeq 0$.

examples

- $\sum_{i=1}^m \log g_i(x)$ is concave if g_i are concave and positive
- $\log \sum_{i=1}^m \exp g_i(x)$ is convex if g_i are convex

Minimization

if $f(x, y)$ is convex in (x, y) and C is a convex set, then

jointly convex in x,y

diff reqs than max

$$g(x) = \inf_{y \in C} f(x, y)$$

is convex

examples

- $f(x, y) = x^T Ax + 2x^T By + y^T Cy$ with • extended quad func.

$$\underbrace{\begin{bmatrix} A & B \\ B^T & C \end{bmatrix}}_{=M} \succeq 0, \qquad C \succ 0$$

• f convex in x if $A \succeq 0$. • in y if $C \succeq 0$.

minimizing over y gives $g(x) = \inf_y f(x, y) = x^T (A - BC^{-1}B^T)x$

g is convex, hence Schur complement $A - BC^{-1}B^T \succeq 0$

Schur complement of M in C

- distance to a set: $\mathbf{dist}(x, S) = \inf_{y \in S} \|x - y\|$ is convex if S is convex
  - To prove using minimization, would have to show $\|x-y\|$ is jointly convex in (x,y). → $\|x-y\|$ is a convex func. of a linear func of $x+y$.

• Reminder that linear ⊂ affine ⊂ convex

Perspective

the **perspective** of a function $f : \mathbf{R}^n \to \mathbf{R}$ is the function $g : \mathbf{R}^n \times \mathbf{R} \to \mathbf{R}$,

another var. t

$$g(x, t) = tf(x/t), \qquad \mathbf{dom}\, g = \{(x, t) \mid x/t \in \mathbf{dom}\, f,\ t > 0\}$$

g is convex if f is convex

examples

- $f(x) = x^T x$ is convex; hence $g(x, t) = x^T x/t$ is convex for $t > 0$ (quad over linear)
- negative logarithm $f(x) = -\log x$ is convex; hence relative entropy $g(x, t) = t \log t - t \log x$ is convex on $\mathbf{R}^2_{++}$
- if f is convex, then (affine, linear fractional)

$$g(x) = (c^T x + d) f\left((Ax + b)/(c^T x + d)\right)$$

is convex on $\{x \mid c^T x + d > 0,\ (Ax + b)/(c^T x + d) \in \mathbf{dom}\, f\}$

internal rate of return

- cash flow $x = (x_0, \ldots, x_n)$; x_i is payment in period i (to us if $x_i > 0$)
- we assume $x_0 < 0$ and $x_0 + x_1 + \cdots + x_n > 0$
- present value of cash flow x, for interest rate r:

$$\mathrm{PV}(x, r) = \sum_{i=0}^{n} (1+r)^{-i} x_i$$

very complicated to max IRR by looking at roots of PV(x,r).

- internal rate of return is smallest interest rate for which $\mathrm{PV}(x, r) = 0$:

• means cash flow equiv. to if money was invested at interest rate = IRR

$$\mathrm{IRR}(x) = \inf\{r \ge 0 \mid \mathrm{PV}(x, r) = 0\}$$

IRR is quasiconcave: superlevel set is intersection of open halfspaces

• Interestingly, the thing we want to do is what we can do (max).

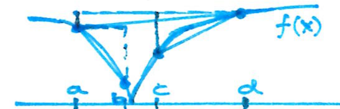
$$\mathrm{IRR}(x) \ge R \iff \sum_{i=0}^{n} (1+r)^{-i} x_i > 0 \text{ for } 0 \le r < R$$

• Can min. convex + quasiconvex

• Can max concave + quasiconcave

• Means PV > 0 for every r below R. So, PV = 0 at r ≥ R.

Properties



chord ~~doesn't~~ not above graph for [a,b] or [c,d]. But, max{f(x), f(y)} is above.

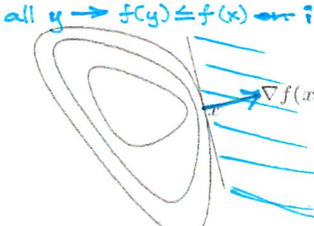
modified Jensen inequality: for quasiconvex f

$$0 \le \theta \le 1 \implies f(\theta x + (1-\theta) y) \le \max\{f(x), f(y)\}$$

first-order condition: differentiable f with cvx domain is quasiconvex $\boxed{\text{iff}}$

$$f(y) \le f(x) \implies \nabla f(x)^T (y - x) \le 0$$

• i.e., all y → f(y) ≤ f(x) ~~are~~ in ↓ halfspace determined by hyperplane det. by ∇f(x)



→ ∇f(x) gives info on how to min. quasiconvex func.

sums of quasiconvex functions are not necessarily quasiconvex

Log-concave and log-convex functions

a positive function f is log-concave if $\log f$ is concave:

• applications show up log-concave much more often.

$$f(\theta x + (1-\theta) y) \ge f(x)^{\theta} f(y)^{1-\theta} \quad \text{for } 0 \le \theta \le 1$$

• Or $\log f(\theta x + (1-\theta)y) \ge \theta \log f(x) + (1-\theta) \log f(y)$

f is log-convex if $\log f$ is convex

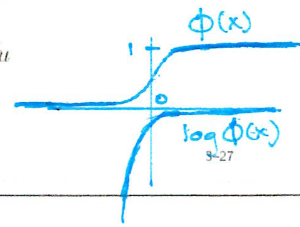
- powers: x^a on $\mathbf{R}_{++}$ is log-convex for $a \le 0$, log-concave for $a \ge 0$
- many common probability densities are log-concave, *e.g.*, normal:

•almost all families log-concave

$$f(x) = \frac{1}{\sqrt{(2\pi)^n \det \Sigma}} e^{\underbrace{-\frac{1}{2}(x-\bar{x})^T \Sigma^{-1} (x - \bar{x})}_{\text{negative quadratic}}}$$

- cumulative Gaussian distribution function Φ is log-concave

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-u^2/2} \, du$$



• domain is ~~$\mathbf{R}_+$~~ $\subseteq \mathbf{R}_+$

• log concave → quasiconcave bc log is monotone

Properties of log-concave functions

- twice differentiable f with convex domain is log-concave if and only if

$$f(x) \nabla^2 f(x) \preceq \nabla f(x) \nabla f(x)^T$$

— dyad and actually a projection mat.
→ rank 1 + PSD
→ $\nabla^2 f(x)$ can have at most one + eigenval.

for all $x \in \operatorname{dom} f$

- product of log-concave functions is log-concave
- sum of log-concave functions is not always log-concave

• gaussian mixture → bimodal but log concave is quasi-concave, so can't have bimodal.

- integration: if $f : \mathbf{R}^n \times \mathbf{R}^m \to \mathbf{R}$ is log-concave, then

$$g(x) = \int f(x, y) \, dy$$

∫ over one variable

is log-concave (not easy to show)

consequences of integration property

- convolution $f * g$ of log-concave functions f, g is log-concave

$$(f * g)(x) = \int f(x-y)g(y)dy$$

- if $C \subseteq \mathbf{R}^n$ convex and y is a random variable with log-concave pdf then

$$f(x) = \mathbf{prob}(x + y \in C)$$

  is log-concave

  proof: write $f(x)$ as integral of product of log-concave functions

$$f(x) = \int g(x+y)p(y)\,dy, \qquad g(u) = \begin{cases} 1 & u \in C \\ 0 & u \notin C, \end{cases}$$

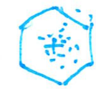
way to rewrite probabilities as indicator func. of set membership.

$$\log(g(u)) = \begin{cases} 0 & u \in C \\ -\infty & u \notin C \end{cases}$$

p is pdf of y

$\rightarrow$ log concave bc C is convex

example: yield function



$$Y(x) = \mathbf{prob}(x + w \in S)$$

- If we assume w log concave, then we can optimize $Y(x)$ for x (x that maximizes $Y(x)$)

- $x \in \mathbf{R}^n$: nominal parameter values for product
- $w \in \mathbf{R}^n$: random variations of parameters in manufactured product
- S: set of acceptable values

if S is convex and w has a log-concave pdf, then

- Y is log-concave
- yield regions $\{x \mid Y(x) \geq \alpha\}$ are convex

Convexity with respect to generalized inequalities

$f : \mathbf{R}^n \to \mathbf{R}^m$ is K-convex if $\mathbf{dom}\, f$ is convex and

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

for $x, y \in \mathbf{dom}\, f$, $0 \leq \theta \leq 1$

example $f : \mathbf{S}^m \to \mathbf{S}^m$, $f(X) = X^2$ is $\mathbf{S}^m_+$-convex

proof: for fixed $z \in \mathbf{R}^m$, $z^T X^2 z = \|Xz\|_2^2$ is convex in X, *i.e.*,

$$z^T(\theta X + (1-\theta)Y)^2 z \leq \theta z^T X^2 z + (1-\theta) z^T Y^2 z$$

for $X, Y \in \mathbf{S}^m$, $0 \leq \theta \leq 1$

therefore $(\theta X + (1-\theta)Y)^2 \preceq \theta X^2 + (1-\theta)Y^2$

- Usually arises for only matrix convexity (i.e., inequality holds in terms of matrices)