Convex Optimization:
CVX 101

Ian Quah (itq)

November 7, 2017

Week 3: Convex Functions

Basic Properties and Examples

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

$$f(\theta x + (1-\theta)y \leq \theta f(x) + (1-\theta)f(y))\ \forall x, y \in \mathbf{dom} f, 0 \leq \theta \leq 1$$

  Aka, if you draw a bowl, with (x, f(x)) on the left side and (y, f(y)) on the right side, the line connecting the two (the chord), is such that the chord is above the graph

- f is concave if -f is convex
- f is strictly convex if **dom** f is convex and

$$f(\theta x + (1-\theta)y < \theta f(x) + (1-\theta)f(y)) \text{ for } x, y \in \mathbf{dom} f, x \neq y, 0 < \theta < 1$$

- Convex examples on $\boldsymbol{R}$
  - affine: ax + b on $\boldsymbol{R}\ \forall a, b \in \boldsymbol{R}$
  - exponential: e^{ax}, for any a $\in \boldsymbol{R}$
  - powers: x^α on $\boldsymbol{R}_{++}$, for $\alpha \geq 1$ or $\alpha \leq 0$
  - powers of absolute value: $|x|^p$ on $\boldsymbol{R}$, for $p \geq 1$
  - Negative entropy: $xlogx$ on $\boldsymbol{R}_{++}$
- Concave examples on $\boldsymbol{R}$
  - affine: ax + b on $\boldsymbol{R} \forall a, b \in \boldsymbol{R}$
  - powers: x^α on $\boldsymbol{R}_{++}$, for $0 \leq \alpha \leq 1$
  - logarithm: $logx$ on $\boldsymbol{R}_{++}$
- Basically if you plot it you have a non-negative curvature (curves up)
- Examples on $\boldsymbol{R}^n$ and $\boldsymbol{R}^{mxn}$

  Affine functions are convex and concave; all norms are convex

  Examples on $\mathbf{R}^n$

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

  Examples on $\mathbf{R}^{mxn}$

  - affine function:

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

  - spectral (maximum singular value) norms

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

- Convexity checking: restricting the convex function to a line
  - f: $\boldsymbol{R}^n \to \boldsymbol{R}$ is convex iff the fn g: $\boldsymbol{R} \to \boldsymbol{R}$

$$g(t) = f(x + tv), \ \textbf{dom } g = \{t | x + tv \in \ \textbf{dom } f\}$$

    - is convex (in t) for any x $\in$ **dom** $f, v \in \boldsymbol{R^n}$
    - can check convexity of f by checking convexity of functions of one variable
    - The intuition here is that a function is convex iff when restricted to all lines it is convex
  - Example: f: $\boldsymbol{S}^n \to \boldsymbol{R}$ with $f(X) = logdetX, \ \textbf{dom } f = \boldsymbol{S}^n_{++}$

$$\begin{aligned} g(t) &= logdet(X + tV) && (X + tV = X^{\frac{1}{2}}(I + X^{-\frac{1}{2}} V X^{-\frac{1}{2}}) X^{\frac{1}{2}}) \\ &= logdetX + logdet(I + tX^{-\frac{1}{2}} V X^{-\frac{1}{2}}) && \text{(log det X constant, } \neg \text{ interesting)} \\ &= logdetX + \sum_{i=1}^{n} log(1 + t\lambda_i) && \text{(RHS is concave)} \end{aligned}$$

  where λ_i are the eigenvalues of $X^{-\frac{1}{2}} V X^{-\frac{1}{2}}$
  g is concave in t (for any choice of X $\succ$0, V); hence f is concave

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

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

  often simplifies notation; e.g

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

  (as an inequality in $\boldsymbol{R} \cup \{\infty\}$), means the same as the two conditions
  - **dom** f is convex
  - for x, y $\in$ **dom** f,

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

Conditions for convexity:
First-Order Condition

- f is differentiable if **dom** f is open and the gradient

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

  , a column vector

  exists at each x $\in$ **dom**f

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

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

  RHS: first order taylor expansion of f at the point x

  first-order approximation of f is a global underestimator

Second-Order Condition

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

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

  exists at each x $\in$ **dom** f

- **2nd-order condition:** for twice differentiable f with convex domain is convex iff
  - f is convex iff

$$\nabla^2 f(x) \succeq 0 \text{ for all x } \in \textbf{ dom } f, \text{the resulting matrix is PSD}$$

  - if $\nabla^2 f(x) \succeq 0$ for all x $\in$ **dom** f, then f is strictly convex

Examples

- quadratic functions: f(x) = $\frac{1}{2}x^T Px + q^T x + r$ with P $\in \boldsymbol{S}^n$

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

  convex if P $\succeq$0

- 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)

- Quadratic-over-linear: f(x, y) = $\frac{x^2}{y}$

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

  convex for y > 0

  - if you plot this and rotate it you get the lorentz cone. It also looks like the front of a ship
  - Is rank 1, and this tells us that at any point it is curving in one direction but not in the other.

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

$$\nabla^2 f(x) = \frac{1}{\mathbf{1}^T z} diag(z) - \frac{zz^T}{(\mathbf{1}^T z)^2}, (z_k = expx_k)$$

  note: LHS is PSD

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

$$\text{since } (\sum_k v_k z_k)^2 \leq (\sum_k z_k v_k^2)(\sum_k z_k)$$

  (from Cauchy-Schwarz inequality)

- geometric mean: f(x) = $(\prod_{k=1}^n x_k)^{\frac{1}{n}}$ on $\boldsymbol{R}^n_{++}$ is concave and follows a similar proof as log-sum-exp

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

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

  sublevel sets of convex functions are convex (converse is false)

- Epigraph of f: $\boldsymbol{R}^n \to \boldsymbol{R}$

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

  - function f is convex IFF **epi** f is a convex set
  - Describes the connection between a convex function and a convex set

- Jensen's inequality
  - Basic inequality: if f is convex, then for $0 \leq \theta \leq 1$,
  $$f(\theta x + (1-\theta)y) \leq \theta f(x) + (1-\theta)f(y)$$
  - Extension: if f is convex, then
  $$f(\mathbf{E}z) \leq \mathbf{E}f(z)$$
  for any random variable z
  - basic inequality is special case with discrete distribution
  prob(z = x) = θ, prob(z = y) = 1 - θ

Operations that preserve Convexity

Practical methods for establishing convexity

- verify defn': often simplified by restricting to a line
- for twice differentiable functions, show ∇^2f(x) $\succeq 0$
- Show that f is obtained from simple convex fns' by operations that themselves preserve convexity
  - 1) non-negative weighted sum, and composition with affine functions
    - nonnegative multiple: αf is convex if f is convex, $\alpha \geq 0$
    - sum: f_1 + f_2 convex, if f_1 and f_2 convex (extends to infinite sums, integrals)
    - f(Ax + b) is convex if f is convex
    - log barrier for linear inequalities:

$$f(x) = -\sum_{i=1}^{m} log(b_i - a_i^T x),\ \textbf{dom}\ f = \{x | a_i^T x < b_i, i = 1, ..., m\}$$

    - (any) norm of affine function: f(x) = $||Ax + b||$
  - 2) Pointwise Max
    - if f_1,...,f_m are convex, then f(x) = max(f_1(x), ..., f_m(x)) is convex

      Intuition: draw multiple bowls and take the intersection
    - piecewise-linear function: f(x) = $\max_{i=1,...,m}(a_i^T x + b_i)$ is convex
    - sum of r largest components of x $\in \boldsymbol{R}^n$:

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

      is convex ($x_{[1]}$ is the ith largest component of x)
      proof:

$$f(x) = max\{x_{i1} + x_{i2} + ... + x_{ir} | 1 \leq i_1 < i_2 < ... < i_r \leq n\}$$

  - 3) 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
        y^Tx is linear, and the supremum of a set of functions is itself convex
      - distance to farthest point in a set C:

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

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

$$\lambda_{max}(X) = sup_{||y||_2 = 1} y^T X y$$

Practical methods for establishing convexity

- Show that f is obtained from simple convex fns' by operations that themselves preserve convexity
    - 4) Composition with scalar functions
        * Composition of g: $\boldsymbol{R}^n \to \boldsymbol{R}$ and $h : \boldsymbol{R} \to \boldsymbol{R}$

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

$$\text{f is convex if } \begin{cases} \text{g convex, h convex, } \tilde{h} \text{ nondecreasing} \\ \text{g concave, h convex, } \tilde{h} \text{ nonincreasing} \end{cases}$$

        * proof