Convex Functions (cont.)

Optimization - 10725
Carlos Guestrin
Carnegie Mellon University
March 3rd, 2008

Function $f: \mathbb{R}^n \rightarrow \mathbb{R}$ is convex if

- Domain is convex
- $\forall x, y \in \text{dom } f$, $\theta \in [0, 1]$,
  $$f(\theta x + (1-\theta)y) \leq \theta f(x) + (1-\theta)f(y)$$

Generalization: Jensen's inequality:

$$f(\mathbb{E}[x]) \leq \mathbb{E}[f(x)]$$

Useful in ML

Strictly convex function:

$$\forall x, y \in \text{dom } f, \theta \in (0, 1), x \neq y, \\theta f(x) + (1-\theta)f(y) < f(\theta x + (1-\theta)y)$$
**Concave functions**

- Function $f$ is concave if
  - $\text{dom} f$ is convex
  - $-f$ is convex

$$f(\omega x + (1-\omega)y) \geq \omega f(x) + (1-\omega)f(y)$$

- Strictly concave:
  - $-f$ is strictly convex

We will be able to optimize:

$$\min_x \begin{cases} f(x) \text{ convex} \quad x \in C \text{ convex set} \end{cases}$$

$$\max_x \begin{cases} f(x) \text{ concave} \quad x \in C \text{ convex set} \end{cases}$$

---

**Proving convexity for a very simple example**

- $f(x) = x^2$

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

$$(\omega x + (1-\omega)y)^2 \leq \omega x^2 + (1-\omega)y^2$$

$$\omega x^2 + \omega x((\omega-1)x + (1-\omega)y) + (1-\omega)y^2 + (1-\omega)y(\omega x - \omega y)$$

$$\omega^2 x^2 + (1-\omega)y^2 + \omega(1-\omega)(x-y)(y-x) \leq 0$$

$$\omega f(x) + (1-\omega)f(y) + 10 \leq 0$$

Done!!

Boring!!

Better way?

Please...
First order condition

- If \( f \) is differentiable in all \( \text{dom} \ f \)

- Then \( f \) convex if and only if \( \text{dom} \ f \) is convex and

Second order condition (1D \( f \))

- If \( f \) is twice differentiable in \( \text{dom} \ f \)

- Then \( f \) convex if and only if \( \text{dom} \ f \) is convex and

- Note 1: Strictly convex if:

- Note 2: \( \text{dom} \ f \) must be convex
  - \( f(x) = 1/x^2 \)
  - \( \text{dom} f = \{x \in \mathbb{R} \mid x \neq 0\} \)
Second order condition (general case)

- If \( f \) is twice differentiable in \( \text{dom} \ f \)

- Then \( f \) convex if and only if \( \text{dom} \ f \) is convex and

- Note 1: Strictly convex if:

Quadratic programming

- \( f(x) = (1/2) x^T A x + b^T x + c \)

- Convex if:
  - Strictly convex if:

- Concave if:
  - Strictly concave if:
Simple examples

- Exponentiation: $e^{ax}$
  - Convex on $\mathbb{R}$, any $a \in \mathbb{R}$

- Powers: $x^a$ on $\mathbb{R}_+$
  - Convex for $a \leq 0$ or $a \geq 1$
  - Concave for $0 < a < 1$

- Logarithm: $\log x$
  - Concave on $\mathbb{R}_+$

- Entropy: $-x \log x$
  - Concave on $\mathbb{R}_+$
  - $(0 \log 0 = 0)$

A few important examples for ML

- Every norm on $\mathbb{R}^n$ is convex

- Log-sum-exp:
  - Convex in $\mathbb{R}^n$

- Log-det:
  - Convex in $\mathbb{S}^{n \times n}_+$
Extended-value extensions

- Convex function $f$ over convex $\text{dom } f$
- Extended-value extension:
  - Still convex:
  - For concave functions
  - Very nice for notation, e.g.,
    - Minimization:
      - Sum:
        - $f_1$ over convex $\text{dom } f_1$
        - $f_2$ over convex $\text{dom } f_2$

Epigraph

- Graph of a function $f: \mathbb{R}^n \rightarrow \mathbb{R}$
  - $\{(x,t) | x \in \text{dom } f, f(x)=t\}$
- Epigraph:
  - $\text{epi } f =$

- Theorem: $f$ is convex if and only if
Support of a convex set and epigraph

- If $f$ is convex & differentiable
  - $f(x) \geq f(x_0) + \nabla f(x_0)^T(x - x_0)$

- For $(x, t) \in \text{epi } f$, $t \geq f(x)$, thus:

- Rewriting:
  $$(x, t) \in \text{epi } f \Rightarrow \begin{bmatrix} \nabla f(x_0) \\ -1 \end{bmatrix}^T \begin{bmatrix} x \\ t \end{bmatrix} - \begin{bmatrix} x_0 \\ f(x_0) \end{bmatrix} \leq 0$$

- Thus, if convex set is defined by epigraph of convex function
  - Obtain support of set by gradient!!
  - If $f$ is not differentiable

Restriction of a convex function to a line

- $f: \mathbb{R}^n \to \mathbb{R}$ convex if and only if $g(t) \ (\mathbb{R} \to \mathbb{R})$ is convex in $t$
  - For all $x_0 \in \text{dom } f$, $v \in \mathbb{R}^n$
    - $\text{dom } g =$

- Can make it much easier to check if $f$ is convex, e.g.,
  - $f(X) = \log \det X$
  - proof in the book...
Operations that preserve convexity

- Many operations preserve convexity
  - Knowing them will make your life much easier when you want to show that something is convex
  - Examples in next few slides

- Simplest: Non-negative weighted sum:
  - If all \( f_i \)'s are convex, then \( f \) is
  - If all \( f_i \)'s are concave, then \( f \) is
  - Example: integral of \( f(x,y) \)

- Affine mapping: \( f: \mathbb{R}^n \rightarrow \mathbb{R} \), \( A \in \mathbb{R}^{n \times m}, b \in \mathbb{R}^m \)
  - \( g(x) = f(Ax+b) \)
  - \( \text{dom } g = \)
    - If \( f \) is convex, then
    - If \( f \) is concave, then

Pointwise maximum and supremum

- If \( f_i \)'s are convex, then

- Piecewise linear convex functions:
  - Fundamental for POMDPs

- For \( x \) in a convex set \( C \), sum of the \( r \) largest elements:
  - Sort \( x \), pick \( r \) largest components, sum them:

- Maximum eigenvalue of symmetric matrix \( X \in \mathbb{R}^{n \times n}, f: \mathbb{R}^{n \times n} \rightarrow \mathbb{R} \)
  - \( f(X) = \)
Pointwise maximum of affine functions: general representation

- We saw: convex set can be written as intersection of (infinitely many) hyperplanes:
  - \( C \) convex, then

- Convex functions can be written as supremum of (infinitely many) lower bounding hyperplanes:
  - \( f \) convex function, then

Discussion on this slide subject to mild conditions on sets and functions, see book

Composition: scalar differentiable, real domain case

- How do I prove convexity of log-sum-exp-positive-weighted-sum-monomials? 😊

- If \( h: \mathbb{R} \to \mathbb{R} \) and \( g: \mathbb{R}^n \to \mathbb{R}^k \), when is \( f(x) = h(g(x)) \) convex (concave)?
  - \( \text{dom} \ f = \{ x \in \text{dom} \ g | g(x) \in \text{dom} \ h \} \)

- Simple case: \( h: \mathbb{R} \to \mathbb{R} \) and \( g: \mathbb{R}^n \to \mathbb{R} \), \( \text{dom} \ g = \text{dom} \ h = \mathbb{R} \), \( g \) and \( h \) differentiable
  - E.g., \( g(x) = x^T \Sigma x \), \( \Sigma \) psd, \( h(y) = e^y \)
  - Second derivative:
    - \( f''(x) = h''(g(x))g'(x)^2 + h'(g(x))g''(x) \)
    - When is \( f''(x) \geq 0 \) (or \( f''(x) \leq 0 \)) for all \( x \)?

- Example of sufficient (but not necessary) conditions:
  - \( f \) convex if \( h \) is convex and nondecreasing and \( g \) is convex
  - \( f \) convex if \( h \) is convex and nonincreasing and \( g \) is concave
  - \( f \) concave if \( h \) is concave and nondecreasing and \( g \) is concave
  - \( f \) concave if \( h \) is concave and nonincreasing and \( g \) is convex
Composition: scalar, general case

- If \( h: \mathbb{R}^k \rightarrow \mathbb{R} \) and \( g: \mathbb{R}^n \rightarrow \mathbb{R}^k \), when is \( f(x) = h(g(x)) \) convex (concave)?
  - \( \text{dom } f = \{ x \in \text{dom } g | g(x) \in \text{dom } h \} \)

- Simple case: \( h: \mathbb{R} \rightarrow \mathbb{R} \) and \( g: \mathbb{R}^n \rightarrow \mathbb{R} \), general domain and non-differentiable
  - Example of sufficient (but not necessary) conditions:
    - \( f \) convex if \( h \) is convex and \( h \) nondecreasing and \( g \) is convex
    - \( f \) convex if \( h \) is convex and \( h \) nonincreasing and \( g \) is concave
    - \( f \) concave if \( h \) is concave and \( h \) nondecreasing and \( g \) is concave
    - \( f \) concave if \( h \) is concave and \( h \) nonincreasing and \( g \) is convex

- Nondecreasing or nonincreasing condition on extend value extension of \( h \) is fundamental
  - Counter example in the book if nondecreasing property holds for \( h \) but not for \( \tilde{h} \), the composition no longer convex
  - If \( h(x)=x^{3/2} \) with \( \text{dom } h = \mathbb{R}^+ \), convex but extension is not nondecreasing
  - If \( h(x)=x^{3/2} \) for \( x \geq 0 \), and \( h(x)=0 \) for \( x<0 \), \( \text{dom } h = \mathbb{R} \), convex and extension is nondecreasing

Vector composition: differentiable

- If \( h: \mathbb{R}^k \rightarrow \mathbb{R} \) and \( g: \mathbb{R}^n \rightarrow \mathbb{R}^k \), when is \( f(x) = h(g(x)) \) convex (concave)?
  - \( \text{dom } f = \{ x \in \text{dom } g | g(x) \in \text{dom } h \} \)
  - Focus on \( f(x) = h(g(x)) = h(g_1(x), g_2(x), \ldots, g_k(x)) \)
  - Second derivative:
    - \( f''(x) = g'(x)^T \nabla^2 h(g(x)) g'(x) + \nabla h(g(x)) g''(x) \)
    - When is \( f''(x) \geq 0 \) (or \( f''(x) \leq 0 \)) for all \( x \)?

- Example of sufficient (but not necessary) conditions:
  - \( f \) convex if \( h \) is convex and nondecreasing in each argument, and \( g_i \) are convex
  - \( f \) convex if \( h \) is convex and nonincreasing in each argument, and \( g_i \) are concave
  - \( f \) concave if \( h \) is concave and nondecreasing in each argument, and \( g_i \) are concave
  - \( f \) concave if \( h \) is concave and nonincreasing in each argument, and \( g_i \) are convex

- Back to log-sum-exp-positive-weighted-sum-monomials
  - \( \text{dom } f = \mathbb{R}^n_+ \), \( c_i > 0 \), \( a_i \geq 1 \)
  - \( \log \text{ sum exp } \text{ convex} \)
Minimization

- If \( f(x,y) \) is convex in \( (x,y) \) and \( C \) is a convex set, then:
  - Norm is convex: \( ||x-y|| \)
    - minimum distance to a set \( C \) is convex:

Perspective function

- If \( f \) is convex (concave), then the perspective of \( f \) is convex (concave):
  - \( t > 0, \ g(x,t) = t f(x/t) \)

- KL divergence:
  - \( f(x) = -\log x \) is convex
  - Take the perspective:
    - Sum over many pairs \( (x_i,t_i) \)