Convex Functions (cont. 2)

Optimization - 10725
Carlos Guestrin
Carnegie Mellon University
March 5th, 2008
Convex Functions

- 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:
  \[ f[E(x)] \leq E[f(x)] \]
  Useful in ML, e.g., EM

- Strictly convex function:
  \( \forall x, y \in \text{dom} f, \ \theta \in (0, 1) \)
  \[ f(\theta x + (1-\theta)y) < \theta f(x) + (1-\theta)f(y) \]
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:
  - \( f = \sum w_i f_i \quad w_i \geq 0 \)
  - If all \( f_i \)'s are convex, then \( f \) is convex
  - If all \( f_i \)'s are concave, then \( f \) is concave
  - Example: integral of \( f(x,y) \)

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

\( \text{e.g., } A, b \text{ come from PCA} \)
Pointwise maximum and supremum

- If $f_i$'s are convex, then $f(x) = \max_i f_i(x)$

- 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:
    $$f(x) = \left\{ \sum_{i=1}^{r} x_{e_i} \mid x_{e_1} > x_{e_2} > \ldots > x_{e_r} \right\}$$

- Maximum eigenvalue of symmetric matrix $X \in \mathbb{R}^{n \times n}$, $f: \mathbb{R}^{n \times n} \rightarrow \mathbb{R}$
  - $f(X) = \max_{\mathbf{A} \subseteq \{1, \ldots, n\}, |\mathbf{A}| = r} \mathbf{A}^T X \mathbf{A}$
  - $f(X) = \max_{\mathbf{V}, \|\mathbf{V}\|_2 = 1} \mathbf{V}^T X \mathbf{V}$
  - $f(X) = \max$ of linear functions of $X$
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
Composition: scalar differentiable, real domain case

- How do I prove convexity of log-sum-exp-positive-weighted-sum-monomials? :)
  - 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 $g: \mathbb{R}^n \rightarrow \mathbb{R}$, $\text{dom } g = \text{dom } h = \mathbb{R}$, $g$ and $h$ differentiable
  - E.g., $g(x) = x^T \sum x$, $\sum \text{ psd}$, $h(y) = e^y$

- 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 \( \tilde{h} \) nondecreasing and \( g \) is convex
    - \( f \) convex if \( h \) is convex and \( \tilde{h} \) nonincreasing and \( g \) is concave
    - \( f \) concave if \( h \) is concave and \( \tilde{h} \) nondecreasing and \( g \) is concave
    - \( f \) concave if \( h \) is concave and \( \tilde{h} \) 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 sum exp 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 $(x_i, t_i)$
Quasiconvex functions

- Unimodal functions are not always convex

- But they are (usually) still easy to optimize: Quasiconvex function:
  - All sublevel sets are convex, for all \( \alpha \in \mathbb{R} \):

- Equivalent definition: max of extremes is higher than function

- Applications include computer vision (geometric reconstruction) [Ke & Kanade ’05]
Log-convex functions

- Function $f: \mathbb{R}^n \rightarrow \mathbb{R}$, with $f(x) > 0$ in all (convex) $\text{dom } f$
  - $f$ log-convex if and only if:

- Or equivalently:

- Examples
  - Gaussian
    $$f(x) = \frac{1}{\sqrt{(2\pi)^n \det \Sigma}} e^{-\frac{1}{2} (x-\mu)^T \Sigma^{-1} (x-\mu)}$$
What should you know: Convex fns

- definition
- showing that a function is convex/concave
  - first principle
  - first and second order condition
  - epigraph
  - operations that preserve convexity
- quasiconvexity
- log-convexity