Convex Functions

Function $f: \mathbb{R}^n \to \mathbb{R}$ 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:
- Useful in ML
  $$E[f(x)] \geq f(E[x])$$

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:
  - $f = \sum w_i f_i \geq 0$
  - 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 $f(x,y)$
    $$g(x) = \int_{y \in C} f(x,y) \, dy$$
    e.g., $A, b$ come from PC

- Affine mapping: $f: \mathbb{R}^n \rightarrow \mathbb{R}$, $A \in \mathbb{R}^{nxm}$, $b \in \mathbb{R}^m$
  - $g(x) = f(Ax+b)$
  - $\text{dom } g = \{ x \mid Ax+b \in \text{dom } f \}$
  - Always convex if $g$ is convex and $\text{dom } f$ convex

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:
    $$f(x) = \sum_{i=1}^{r} x_i$$
    $$\max \left\{ \sum_{i=1}^{r} x_i \right\}$$

- Maximum eigenvalue of symmetric matrix $X \in \mathbb{R}^{nxn}$, $f: \mathbb{R}^{nxn} \rightarrow \mathbb{R}$
  - $f(X) = \max \left\{ \sum_{i=1}^{n} x_i \right\}$
  - Linear function of $X$
  - $\lambda_{\text{max}}$
Pointwise maximum of affine functions: general representation

We saw: convex set can be written as intersection of (infinitely many) hyperplanes:
- \( C \) convex, then
  \[ C = \bigcap_i (a_i^T x + b_i \geq 0) \]

Convex functions can be written as supremum of (infinitely many) lower bounding hyperplanes:
- \( f \) convex function, then
  \[ f(x) = \max_i a_i^T x + b_i \]

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? :)
- \( f(x) = \log \sum_i e^{c_i^T x + d_i} \)
- If \( h: \mathbb{R}^k \to \mathbb{R} \) and \( g: \mathbb{R}^n \to \mathbb{R}^k \), when is \( f(x) = h(g(x)) \) convex (concave)?
  - \( \text{dom } f = \{ x \in \text{dom } g \mid g(x) \in \text{dom } h \} \)

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., \( g(x) = x^T \sum x, \sum \text{psd}, h(y) = e^y \)

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)?

- $\text{dom } f = \{x \in \text{dom } g \mid g(x) \in \text{dom } h\}$

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 $g$ is nondecreasing
  - $f$ convex if $h$ is convex and $g$ is 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 $h(x) = x^p$ with $\text{dom } h = \mathbb{R}^+$, convex but extension is not nondecreasing

  - $\tilde{h}(x) = \begin{cases} x^{1/2} & x > 0 \\ \infty & x < 0 \end{cases}$

- If $h(x) = x^p$ 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)?

- $\text{dom } f = \{x \in \text{dom } g \mid g(x) \in \text{dom } h\}$

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

- $f(x) = \sum_i a_i x_i^p$
- $\text{dom } f = \mathbb{R}_{++}^n$, $c > 0$, $a_i > 1$

- $f(x)$ is nondecreasing in each argument if $f$ is convex
Minimization

- If \( f(x,y) \) is convex in \((x,y)\) and \( C \) is a convex set, then:
  \[
  g(y) = \min_{x \in C} f(x,y)
  \]

- Norm is convex: \( ||x-y|| \)
  \[
  (x-y) = A \begin{bmatrix} x \ y \end{bmatrix}^T
  \]
  - minimum distance to a set \( C \) is convex:
  \[
  d(x, C) = \min_{y \in C} ||x-y||
  \]

Perspective function

- If \( f \) is convex (concave), then the perspective of \( f \) is convex (concave):
  \[ t > 0, \ g(x,t) = t f(x/t) \]
  - computer vision
  - \( f \) convex in \( x \), \( g \) is convex in both \( x, t \)

- KL divergence:
  - \( f(x) = -\log x \) is convex
  - Take the perspective:
    \[
    g(x,t) = -t \log \frac{x}{t}
    \]
  - Sum over many pairs \((x_i, t_i)\)
    \[
    h(x, t) = \sum_i g(x_i, t_i)
    \]
    \[
    KL(p || q) = \sum_x p(x) \log \frac{p(x)}{q(x)}
    \]
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 \( \alpha \in \mathbb{R} \):
    \[
    S_\alpha = \{ x \mid f(x) \leq \alpha \}
    \]
- Equivalent definition: max of extremes is higher than function
  \[
  f(\theta x + (1-\theta)y) \leq \theta f(x) + (1-\theta)f(y)
  \]
  \[
  f(\theta x + (1-\theta)y) \leq \max(f(x), f(y))
  \]
- 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:
    \[
    \log f(\theta x + (1-\theta)y) \leq \theta \log f(x) + (1-\theta) \log f(y)
    \]
  - Or equivalently:
    \[
    f(\theta x + (1-\theta)y) \leq f(x) ^ \theta f(y) ^ (1-\theta)
    \]
- Examples
  - Gaussian
  \[
  f(x) = \frac{1}{\sqrt{(2\pi)^d |\Sigma|}} e^{-\frac{1}{2} (x-M)^T \Sigma^{-1} (x-M)}
  \]
  \[
  \log f(x) = \log \frac{1}{\sqrt{(2\pi)^d |\Sigma|}} - \frac{1}{2} (x-M)^T \Sigma^{-1} (x-M)
  \]
  \[
  \log \text{convex}
  \]
  \[
  \log \text{concave}
  \]
  \[
  \log \text{concave}
  \]

©2008 Carlos Guestrin
11
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