# Convex Functions (cont.)

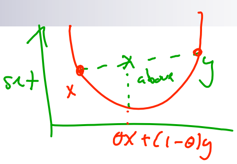
Optimization - 10725
Carlos Guestrin
Carnegie Mellon University

March 3rd, 2008

# Convex Functions

- Function $f:R^n \rightarrow R$ is convex if
  - Domain is convex  dom f convex set
  - $\forall x,y \in \text{dom} f$, $\theta \in [0,1]$




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

- Generalization: Jensen's inequality:

prob dist x ∈ convex domain of f

$$f[E[x]] \leq E_x[f(x)]$$

useful in ML
e.g., EM

- Strictly convex function:

e.g., ∪ , convex non-strict

$\forall x,y \in \text{dom} f$, $\theta \in (0,1)$
$x \neq y$

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

# Concave functions

- Function f is concave if
  - dom f is convex
  - $-f$ is convex

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

- Strictly concave: $-f$ is strictly convex

- We will be able to optimize: "easy"

$\min_x f(x)$ convex
$x \in C$ convex set

$\max_x f(x)$ concave
$x \in C$ convex set



# Proving convexity for a very simple example

- $f(x)=x^2$

$$f(\theta x + (1-\theta)y) \overset{?}{\leq} \theta f(x) + (1-\theta) f(y)$$

$$(\theta x + (1-\theta)y)^2 \overset{?}{\leq} \theta x^2 + (1-\theta)y^2$$

$$\theta x(\theta x \overset{+x-x}{+} (1-\theta)y) + (1-\theta)y(\theta x + (1-\theta)y)$$

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

$$\underbrace{\theta x^2 + (1-\theta)y^2}_{\theta f(x) + (1-\theta)f(y)} + \underbrace{\theta(1-\theta)}_{\geq 0}\underbrace{(x-y)(y-x)}_{\leq 0}$$

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

done!!

boring!!

better way?

Please...

# First order condition

- If f is differentiable in all **dom** f
- Then f convex if and only if **dom** f is convex and




# Second order condition (1D f)

- If f is twice differentiable in **dom** f
- Then f convex if and only if **dom** f is convex and
- Note 1: Strictly convex if:
- Note 2: **dom** f must be convex
  - $f(x)=1/x^2$
  - **dom** f = $\{x \in R | x \neq 0\}$

## Second order condition (general case)

- If f is twice differentiable in **dom** f
- Then f convex if and only if **dom** f is convex and
- Note 1: Strictly convex if:



## Quadratic programming

- $f(x) = (1/2)\ x^TAx + b^Tx + c$
- Convex if:
- Strictly convex if:
- Concave if:
- Strictly concave if:

## Simple examples

- Exponentiation: $e^{ax}$
  - convex on R, any $a \in R$
- Powers: $x^a$ on $R_{++}$
  - Convex for $a \leq 0$ or $a \geq 1$
  - Concave for $0 \leq a \leq 1$
- Logarithm: log x
  - Concave on $R_{++}$
- Entropy: -x log x
  - Concave on $R_+$
  - (0 log 0 = 0)




## A few important examples for ML

- Every norm on $R^n$ is convex
- Log-sum-exp:
  - Convex in $R^n$
- Log-det:
  - Convex in $S^n_{++}$

# Extended-value extensions

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




# Epigraph

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

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

- 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

- $f: R^n \rightarrow R$ convex if and only if g(t) ($R \rightarrow R$) is convex in t
  - For all $x_0 \in$ **dom** f, $v \in R^n$
  -
  - **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: $f:R^n \rightarrow R$, $A \in R^{n \times m}$, $b \in R^m$
  - g(x) = f(Ax+b)
  - **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 $X \in R^{n \times n}$, $f:R^{n \times n} \rightarrow R$
  - 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 $h:R^k \rightarrow R$ and $g:R^n \rightarrow R^k$, when is $f(x) = h(g(x))$ convex (concave)?
  - **dom** f = {x ∈ **dom** g| g(x) ∈ **dom** h}

- Simple case: $h:R \rightarrow R$ and $g:R^n \rightarrow R$, **dom** g = **dom** h = 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 $h:R^k \rightarrow R$ and $g:R^n \rightarrow R^k$, when is $f(x) = h(g(x))$ convex (concave)?
  - **dom** f = {x ∈ **dom** g| g(x) ∈ **dom** h}
- Simple case: $h:R \rightarrow R$ and $g:R^n \rightarrow R$, 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 **dom** h = $R_+$, convex but extension is not nondecreasing
  - If $h(x)=x^{3/2}$ for $x\geq0$, and $h(x)=0$ for $x<0$, **dom** h = R, convex and extension is nondecreasing



# Vector composition: differentiable

- If $h:R^k \rightarrow R$ and $g:R^n \rightarrow R^k$, when is $f(x) = h(g(x))$ convex (concave)?
  - **dom** f = {x ∈ **dom** g| g(x) ∈ **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 $f''(x)\geq0$ (or $f''(x)\leq0$) 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
  -
  - **dom** f = $R^n_{++}$ , $c_i>0$, $a_i\geq1$
  - 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 ($x_i$,$t_i$)