

CO367: Nonlinear Optimization
Lecture 8, Thursday Jan. 31, 2013.

Henry Wolkowicz

Dept. Combinatorics and Optimization, University of Waterloo

February 1, 2013

f is a convex function iff

- 1 (0-order: secant lines lie above the graph)

$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda)f(y), \forall 0 \leq \lambda \leq 1, \forall x, y \in D.$$

iff

- 2 (1-order: tangent planes lie below the graph)

$$\nabla f(x)^T (y - x) \leq f(y) - f(x), \forall x, y \in D$$

iff

- 3 (2-order: Hessians are psd (curvature))

$$\nabla^2 f(x) \succeq 0, \forall x \in D \quad (\text{psd})$$

Definition:

As above, let $f : D \rightarrow \mathbb{R}$, $D \subset \mathbb{R}^n$ open convex set. Then the epigraph of f is

$$\text{epi}(f) = \{(x, r) \in \mathbb{R}^n \times \mathbb{R} : x \in D, f(x) \leq r\}.$$

Theorem

Let $f : D \rightarrow \mathbb{R}$, $D \subset \mathbb{R}^n$ open convex set. Then f is a convex function if, and only if, $\text{epi}(f)$ is a convex set.

The proof is immediate from the [0-order characterization](#), i.e., from the fact that the secant lines lie above the graph.

Convex Function Preserving Operations

Let $h : \mathbb{R} \rightarrow \mathbb{R}$, $g : \mathbb{R}^n \rightarrow \mathbb{R}$ and the composition (domains/differentiability appropriately defined/assumed). And define the composition

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

Then the second derivative

$$f''(x) = h''(g(x))g'(x)^2 + h'(g(x))g''(x).$$

f is convex if either condition holds:

- 1 h is convex and nondecreasing and g is convex
- 2 h is convex and nonincreasing and g is concave

Examples of Composite Convex Functions

- 1 g convex implies $\exp(g(x))$ is convex
- 2 g convex, nonnegative, $p \geq 1$ implies $(g(x))^p$ is convex
- 3 g convex, implies $-\log(-g(x))$ is convex on $\{x : g(x) < 0\}$.

Further Convexity Preserving Operations

Let $f_i : D \rightarrow \mathbb{R}$ be convex functions, D convex set

- 1 set of convex functions on suitable convex domain $D \subseteq \mathbb{R}^n$ forms a convex cone, i.e., closed under addition and nonneg. scalar multipl.: $g(x) := \sum_{i=1}^k \lambda_i f_i(x)$, $\lambda_i \geq 0, \forall i$ (proof: e.g. use 2nd-order Hessian characterization)
- 2 $g(x) := \sup_{i \in I} \{f_i(x)\}$ (proof: use the epigraph characterization and intersection of epigraphs)

Applications of \sup :

- 1 $\lambda_{\max}(A) = \max_{\|x\|=1} x^T A x$ (largest eigenvalue)
- 2 $g(x) := \sup_{y \in C} \|x - y\|$ (distance to furthest point in C)

Arithmetic-Geometric Mean Inequality

Consider $\max\{\sqrt[n]{\prod_{i=1}^n x_i} : \frac{1}{n} \sum_{i=1}^n x_i = 1\}$. We can take logs and scale without changing the optimal \mathbf{x} to get

$$\max\{\sum_{i=1}^n \log x_i : \sum_{i=1}^n x_i = 1\}.$$

We can use Lagrange multipliers (one multiplier) to get the Lagrangian $L(\mathbf{x}, \lambda) = \sum_{i=1}^n \log x_i + \lambda(1 - \sum_{i=1}^n x_i)$. and

$$0 = \nabla L(\mathbf{x}, \lambda) = ((1/x_i) - \lambda),$$

i.e., all x_i are equal. The optimal solution for $\sum_{i=1}^n x_i = 1$ is $x_i = \frac{1}{n}$. Therefore

$$\text{GM} = \sqrt[n]{\prod_{i=1}^n x_i} = \frac{1}{n} = \frac{1}{n} \sum_{i=1}^n x_i = \text{AM}$$

at the maximum for the GM. Conclusion: $\text{GM} \leq \text{AM}$ with equality if, and only if all x_i are equal.

Other Inequalities

Similarly: generalized AGM and other inequalities, e.g., Cauchy-Schwartz and Holder inequalities, can be proved this way.

Problem: Find the open rectangular box with a fixed surface area S_0 that has the largest volume.

Solution

$$\begin{aligned} S_0 &= x_1 x_2 + 2x_1 x_3 + 2x_2 x_3 \\ &= 3 \left(\frac{x_1 x_2 + 2x_1 x_3 + 2x_2 x_3}{3} \right) \\ &\geq 3 \left((x_1 x_2)^{1/3} (2x_1 x_3)^{1/3} (2x_2 x_3)^{1/3} \right) \\ &= 3(4)^{1/3} (x_1^2 x_2^2 x_3^2)^{1/3} \\ &= 3(4)^{1/3} V^{2/3} \end{aligned}$$

and V is max when $x_1 x_2 = 2x_1 x_3 = 2x_2 x_3 = S_0/3$. Yields

$$x_1 = x_2 = \sqrt{\frac{S_0}{3}} \text{ and } x_3 = \frac{1}{2} \frac{S_0}{3}$$

Thanks for your attention!

CO367: Nonlinear Optimization
Lecture 8, Thursday Jan. 31, 2013.

Henry Wolkowicz

Dept. Combinatorics and Optimization, University of Waterloo

February 1, 2013