

Convex Optimization: Part I

CMPT 419/983

23/09/2018

Textbook

- S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2008.

Outline

- Optimization program
 - Examples and classes
- Convex optimization
 - Convex functions
 - Optimality conditions
- Numerical solutions

Optimization Program: Terminology

minimize $f(x)$

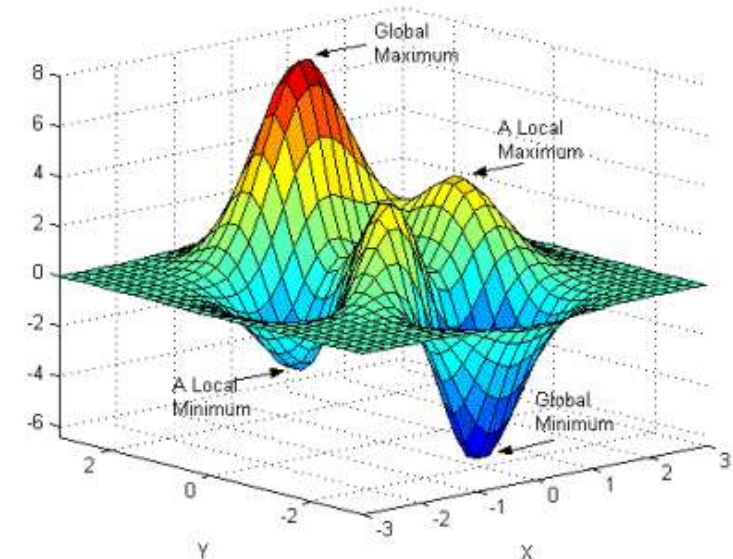
subject to $g_i(x) \leq 0, i = 1, \dots, n$
 $h_j(x) = 0, j = 1, \dots, m$

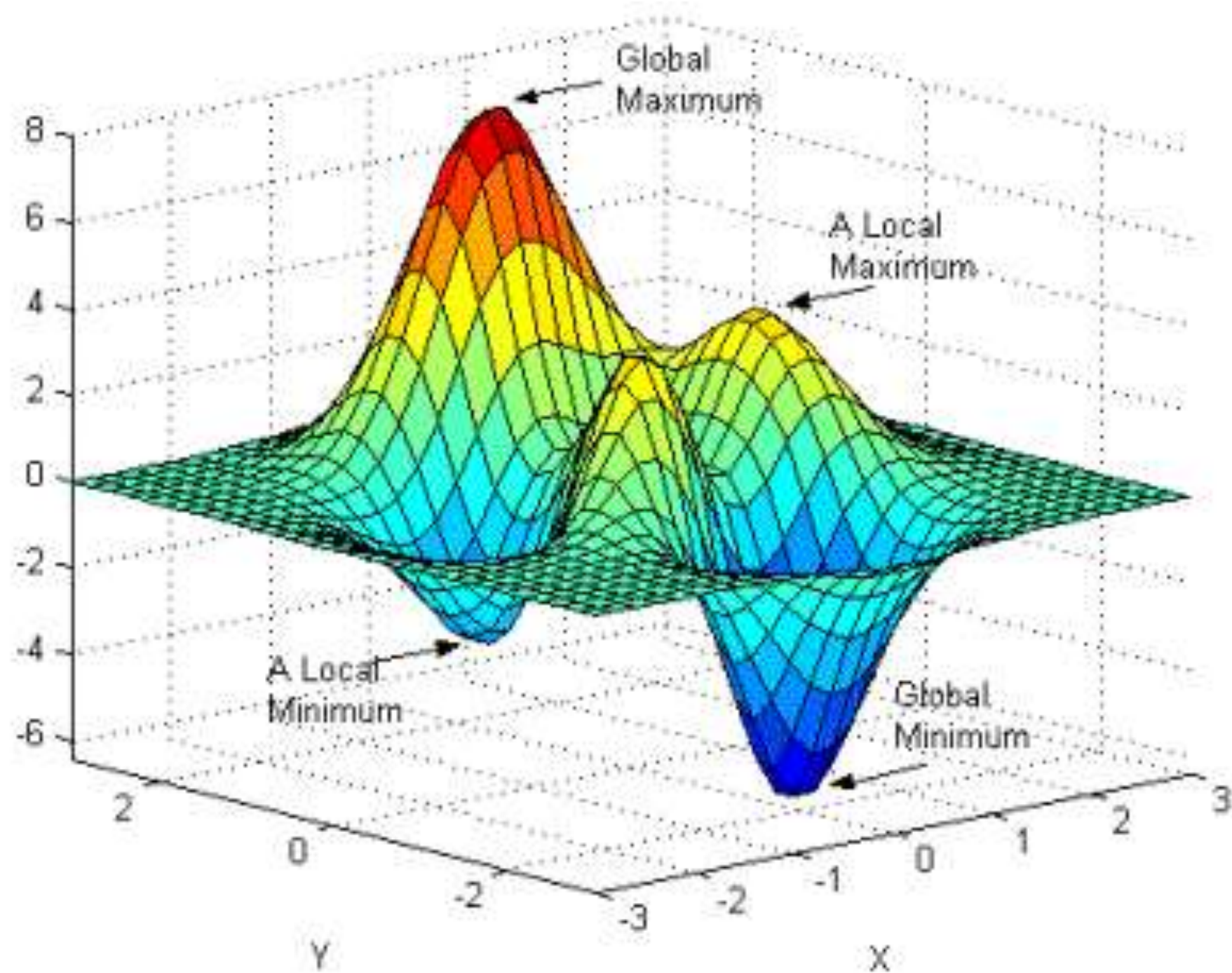
Objective function

Inequality constraints

Equality constraints

- For now, assume f, g_i, h_j are twice differentiable
- Look for an **optimal solution**, the vector x^*
 - **Locally optimal**: x^* is a local minimum of $f(x)$
 - **Globally optimal**: x^* is a global minimum of $f(x)$





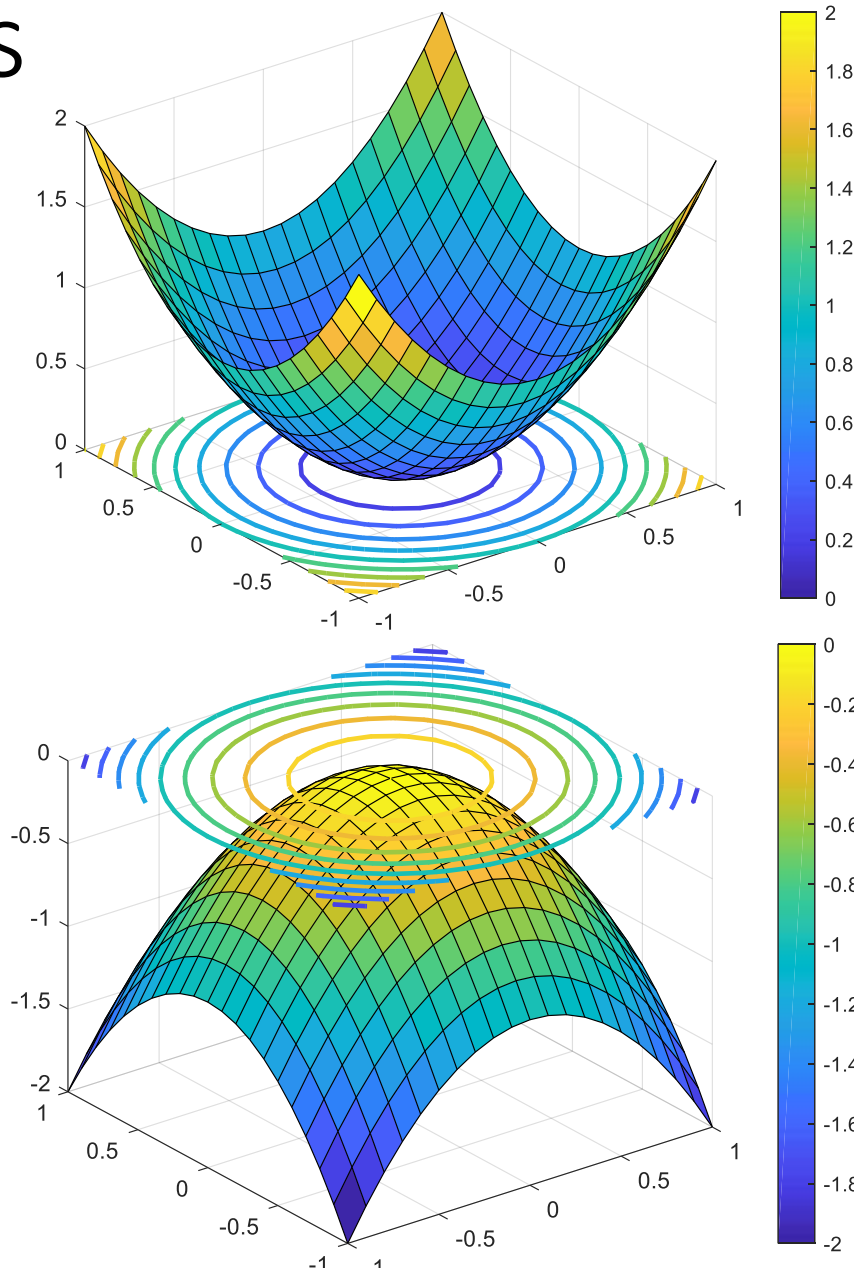
Optimization Program: Examples

$$\begin{array}{ll}\text{minimize} & f(x) \\ \text{subject to} & g_i(x) \leq 0, i = 1, \dots, n \\ & h_j(x) = 0, j = 1, \dots, m\end{array}$$

- Applications: Portfolio management

$$\begin{array}{ll}\text{maximize} & \text{Expected profit} \\ \text{subject to} & \text{Maximum budget} \\ & \text{Maximum acceptable risk}\end{array}$$

$$\min f(x) = -\max\{-f(x)\}$$



Optimization Program: Examples

$$\begin{array}{ll}\text{minimize} & f(x) \\ \text{subject to} & g_i(x) \leq 0, i = 1, \dots, n \\ & h_j(x) = 0, j = 1, \dots, m\end{array}$$

- Applications: Portfolio management

$$\begin{array}{ll}\text{minimize} & \text{Overall risk} \\ \text{subject to} & \text{Maximum budget} \\ & \text{Minimum acceptable expected profit}\end{array}$$

Constraints vs. objectives

- Sometimes constraints can be “moved” to the objective as a “penalty”

Optimization Program: Examples

$$\begin{array}{ll}\text{minimize} & f(x) \\ \text{subject to} & g_i(x) \leq 0, i = 1, \dots, n \\ & h_j(x) = 0, j = 1, \dots, m\end{array}$$

- Applications: Building heating, ventilation, and air conditioning

$$\begin{array}{ll}\text{minimize} & \text{Energy consumption} \\ \text{subject to} & \text{Acceptable temperature range by location} \\ & \text{Acceptable noise level} \\ & \text{Internal and external heat transfer}\end{array}$$

Optimization Program: Examples

$$\begin{aligned} &\text{minimize} && f(x) \\ &\text{subject to} && g_i(x) \leq 0, i = 1, \dots, n \\ & && h_j(x) = 0, j = 1, \dots, m \end{aligned}$$

- Applications: Robotic trajectory planning

$$\begin{aligned} &\text{minimize} && \text{Fuel consumption} \\ &\text{subject to} && \text{Goal reaching} \\ & && \text{System dynamics} \\ & && \text{Collision avoidance} \end{aligned}$$

Optimization Program: Examples

$$\begin{aligned} &\text{minimize} && f(x) \\ &\text{subject to} && g_i(x) \leq 0, i = 1, \dots, n \\ & && h_j(x) = 0, j = 1, \dots, m \end{aligned}$$

- Applications: Robotic trajectory planning

$$\begin{aligned} &\text{minimize} && \text{Distance to goal} \\ &\text{subject to} && \text{Fuel limitations} \\ & && \text{System dynamics} \\ & && \text{Collision avoidance} \end{aligned}$$

Optimization Program: Examples

$$\begin{aligned} &\text{minimize} && f(x) \\ &\text{subject to} && g_i(x) \leq 0, i = 1, \dots, n \\ & && h_j(x) = 0, j = 1, \dots, m \end{aligned}$$

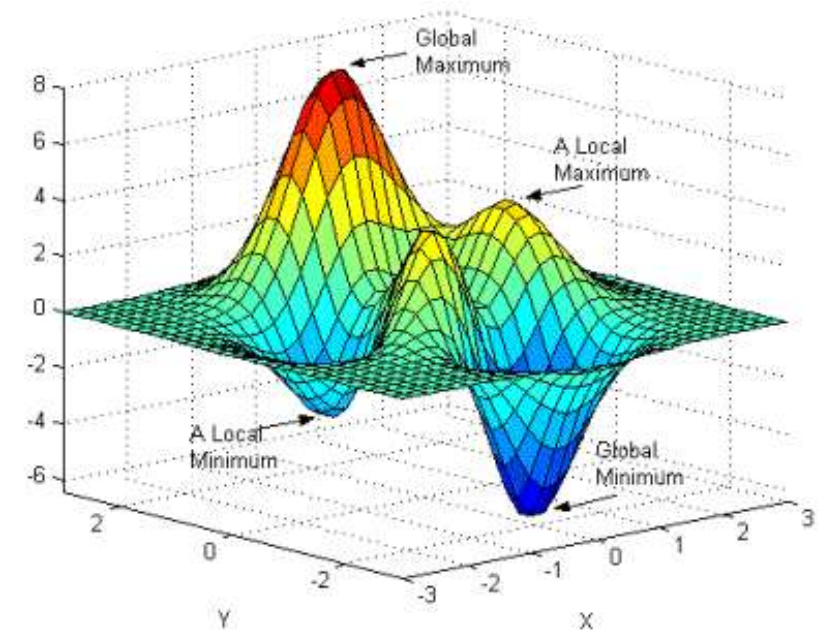
- Applications: Machine learning

$$\begin{aligned} &\text{maximize} && \text{Performance (eg. Accuracy of object recognition)} \\ &\text{subject to} && \text{Problem constraints} \end{aligned}$$

Optimization Program

$$\begin{array}{ll}\text{minimize} & f(x) \\ \text{subject to} & g_i(x) \leq 0, i = 1, \dots, n \\ & h_j(x) = 0, j = 1, \dots, m\end{array}$$

- Very difficult to solve in general
 - Trade-offs to consider: computation time, solution optimality
- Easy cases:
 - Find global optimum for **linear program**: f, g_i, h_j are linear
 - Find global optimum for **convex program**: f, g_i are convex, h_j is linear
 - Find local optimum for **nonlinear program**: f, g_i, h_j are differentiable



Example: Least Squares

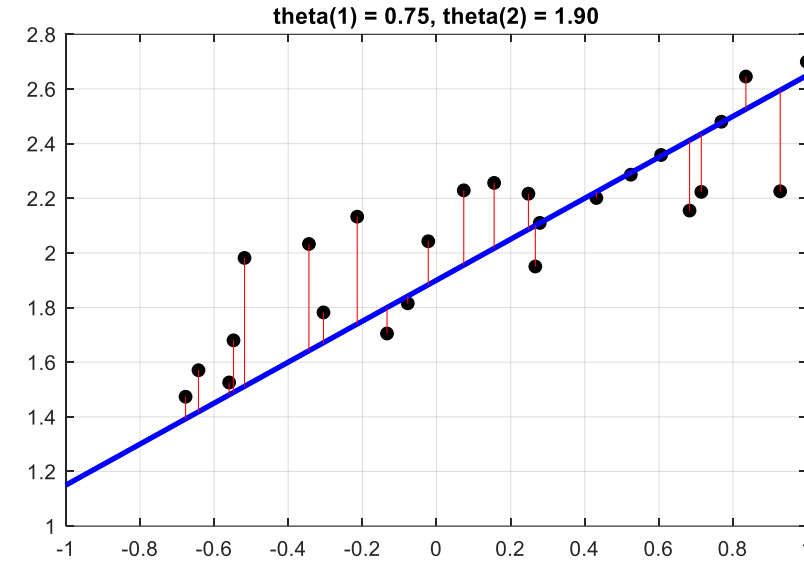
$$\underset{\theta}{\text{minimize}} \|X\theta - Y\|_2^2$$

- Scalar example:

- Data: $\{x_i, y_i\}_{i=1}^n, x_i, y_i \in \mathbb{R}$
- Model: $y = mx + b, m, b \in \mathbb{R}$
- Sum of error of model: $\sum_{i=1}^n (y_i - mx_i - b)^2$
- No constraints: allow *any* m, b

- Error in matrix form: $e_i = y_i - [x_i \quad 1] \begin{bmatrix} m \\ b \end{bmatrix}$

- Stacking the data points:
$$E_i = \underbrace{\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}}_Y - \underbrace{\begin{bmatrix} x_1 & 1 \\ x_2 & 1 \\ \vdots & \vdots \\ x_n & 1 \end{bmatrix}}_X \underbrace{\begin{bmatrix} m \\ b \end{bmatrix}}_\theta$$

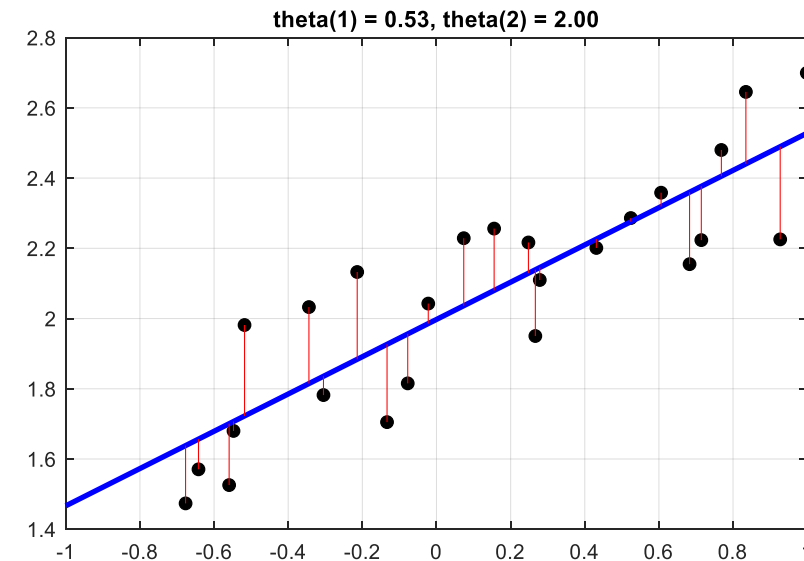
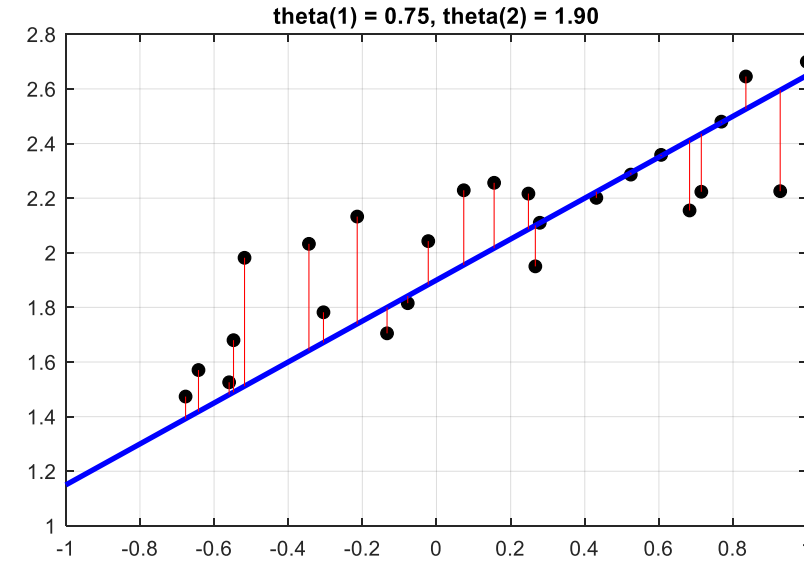


Example: Least Squares

$$\underset{\theta}{\text{minimize}} \|X\theta - Y\|_2^2$$

- Analytic solution available!
 - Objective: $f(\theta) = \|X\theta - Y\|_2^2$, set derivative to zero
 - $f(\theta) = (X\theta - Y)^\top (X\theta - Y)$
 - $f(\theta) = \theta^\top X^\top X\theta - 2Y^\top X\theta + Y^\top Y$

$$\begin{aligned}\frac{\partial f}{\partial \theta} &= 2X^\top X\theta - 2X^\top Y \\ 0 &= 2X^\top X\theta - 2X^\top Y \\ X^\top Y &= X^\top X\theta \\ \theta &= (X^\top X)^{-1} X^\top Y\end{aligned}$$



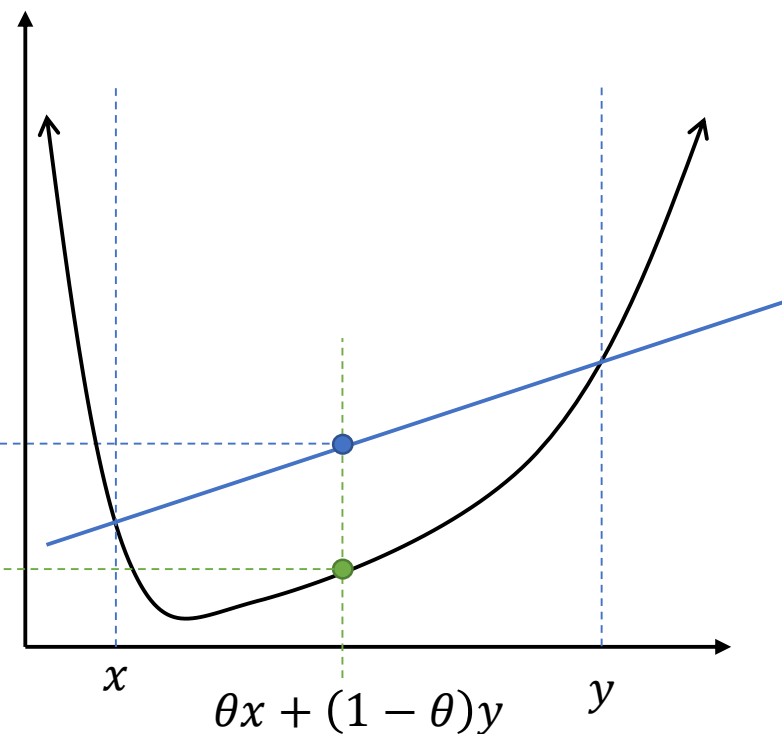
Convex Programs

minimize $f(x)$

subject to $g_i(x) \leq 0, i = 1, \dots, n,$
where $g_i(x)$ are convex
 $h_j^\top x = 0, j = 1, \dots, m$

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

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



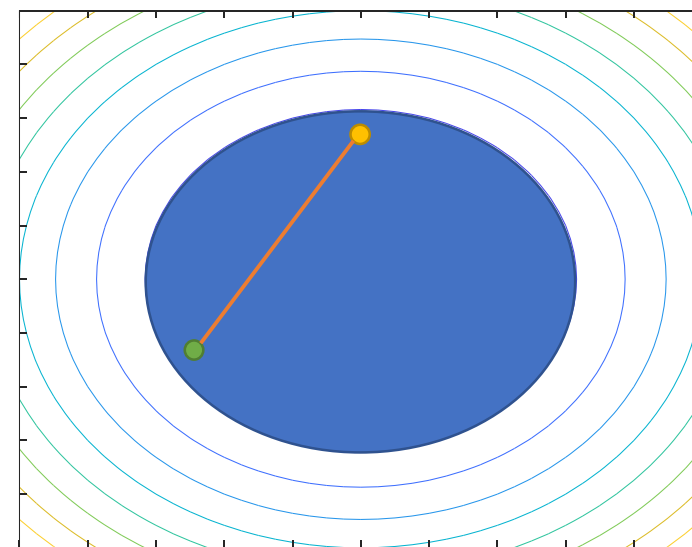
- **Convex function**

$f(\theta x + (1 - \theta)y) \leq \theta f(x) + (1 - \theta)f(y)$ for all $x, y \in \mathbb{R}^n$,
for all $\theta \in [0, 1]$

- Sublevel sets of convex functions, $\{x: f(x) \leq C\}$, are convex

- **Convex shape \mathcal{C} :**

$$x_1, x_2 \in \mathcal{C}, \theta \in [0, 1] \Rightarrow \theta x_1 + (1 - \theta)x_2 \in \mathcal{C}$$



Convex Programs

minimize $f(x)$

subject to $g_i(x) \leq 0, i = 1, \dots, n,$
where $g_i(x)$ are convex
 $h_j^\top x = 0, j = 1, \dots, m$

- **Convex function**

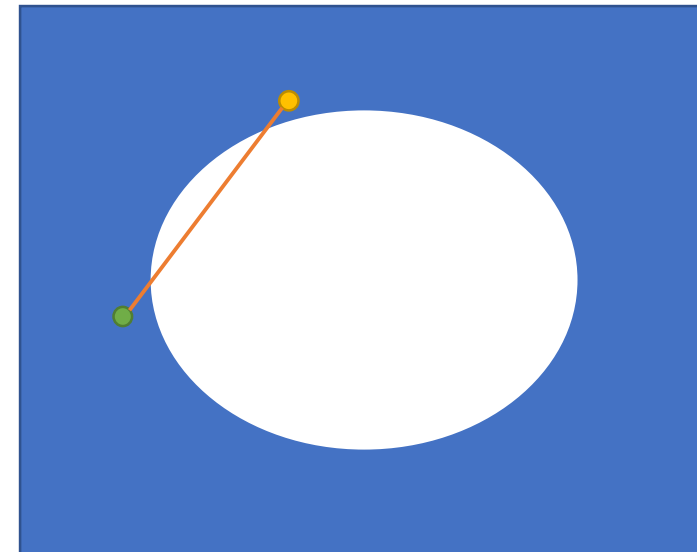
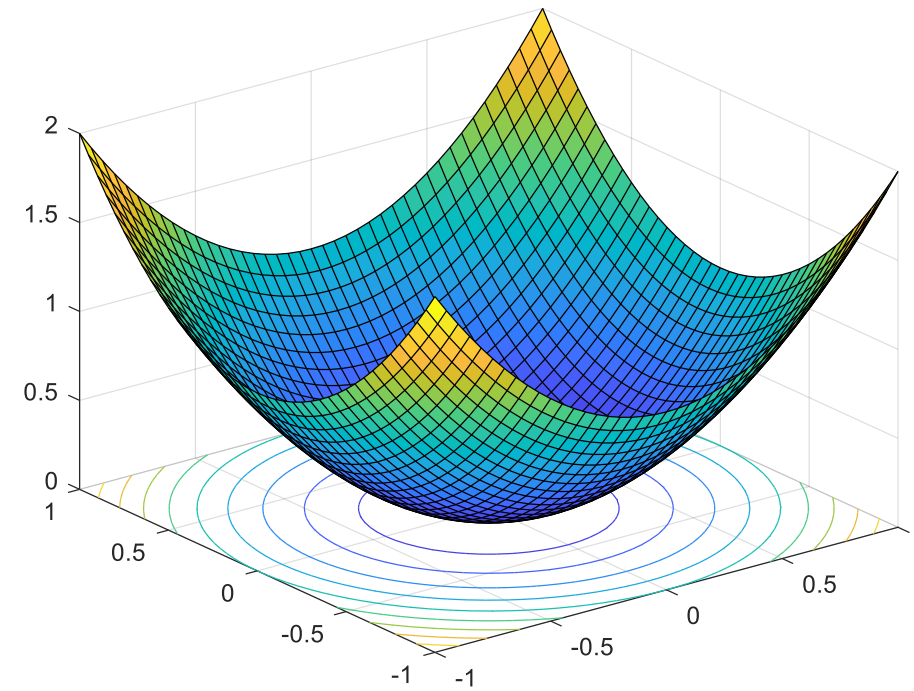
$f(\theta x + (1 - \theta)y) \leq \theta f(x) + (1 - \theta)f(y)$ for all $x, y \in \mathbb{R}^n$,
for all $\theta \in [0, 1]$

- Sublevel sets of convex functions, $\{x: f(x) \leq C\}$, are convex

- **Convex shape \mathcal{C} :**

$x_1, x_2 \in \mathcal{C}, \theta \in [0, 1] \Rightarrow \theta x_1 + (1 - \theta)x_2 \in \mathcal{C}$

- Superlevel sets of convex functions are *not* convex!



Convex Programs

minimize $f(x)$

subject to $g_i(x) \leq 0, i = 1, \dots, n,$
where $g_i(x)$ are convex
 $h_j^\top x = 0, j = 1, \dots, m$

- **Convex function**

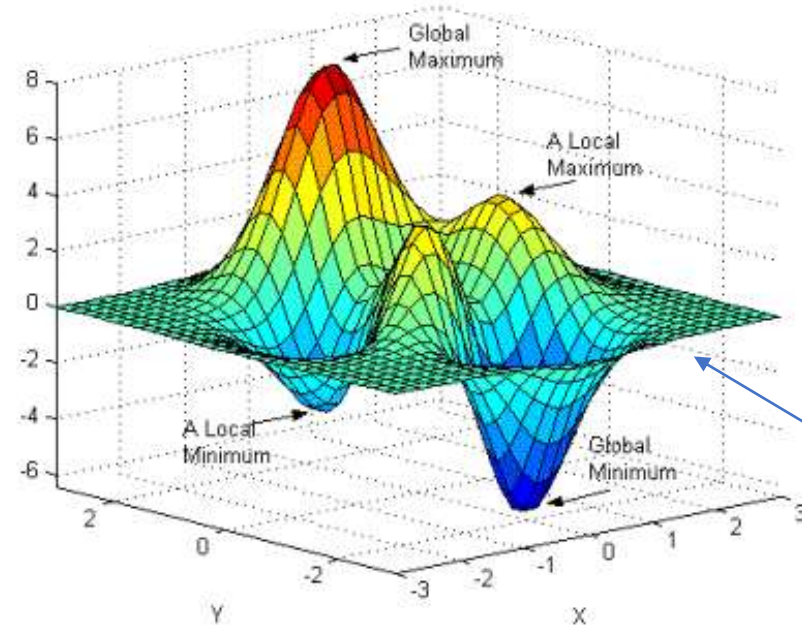
$f(\theta x + (1 - \theta)y) \leq \theta f(x) + (1 - \theta)f(y)$ for all $x, y \in \mathbb{R}^n$,
for all $\theta \in [0,1]$

- Sublevel sets of convex functions, $\{x: f(x) \leq C\}$, are convex

- **Convex shape \mathcal{C} :**

$x_1, x_2 \in \mathcal{C}, \theta \in [0,1] \Rightarrow \theta x_1 + (1 - \theta)x_2 \in \mathcal{C}$

- Superlevel sets of convex functions are *not* convex!



Not convex

Convex Programs

minimize $f(x)$, where f is convex

subject to $g_i(x) \leq 0, i = 1, \dots, n,$
 where $g_i(x)$ are convex
 $h_j^\top x = 0, j = 1, \dots, m$

minimize A convex objective function

subject to Convex inequality constraints
 Linear equality constraints

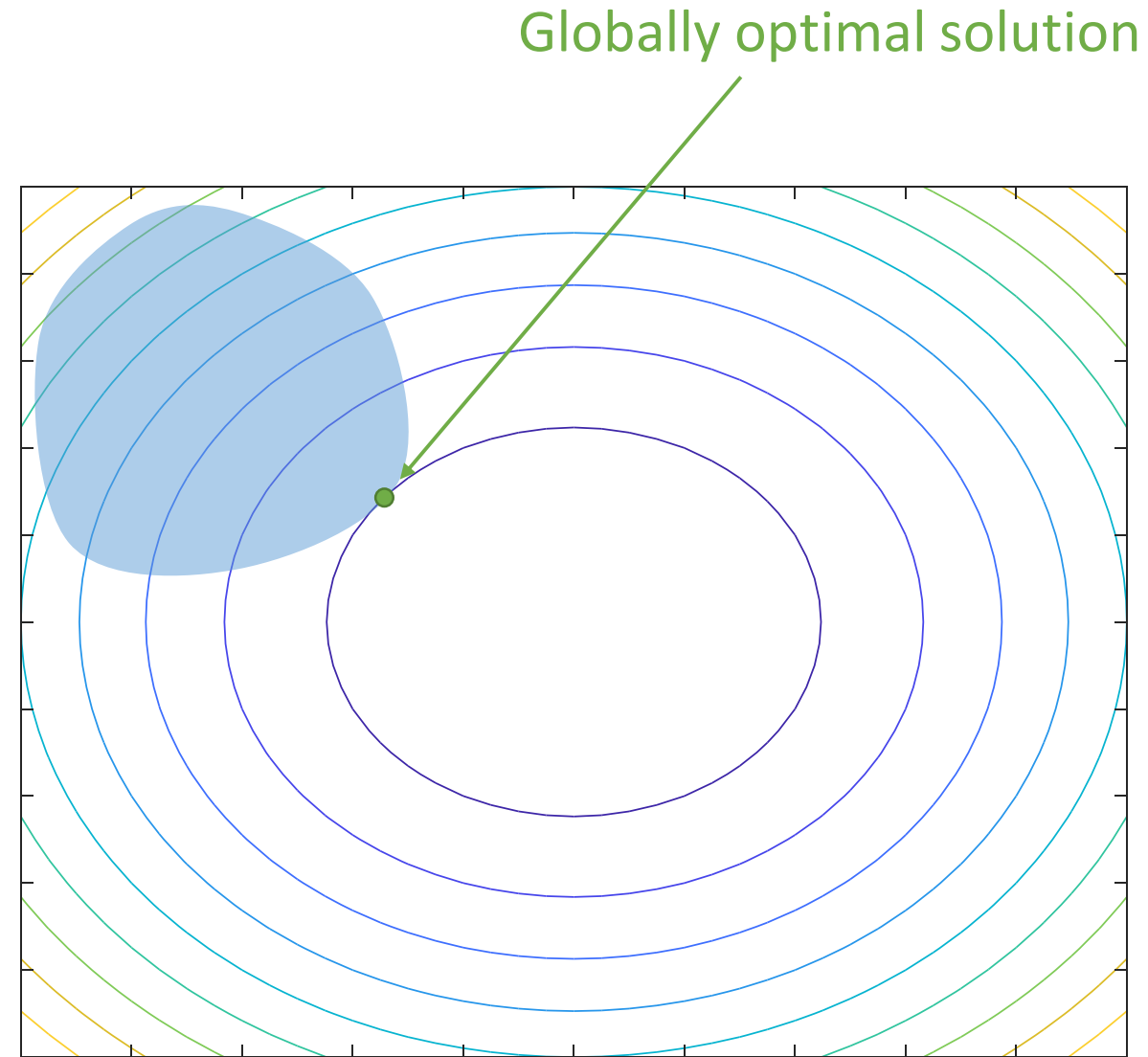
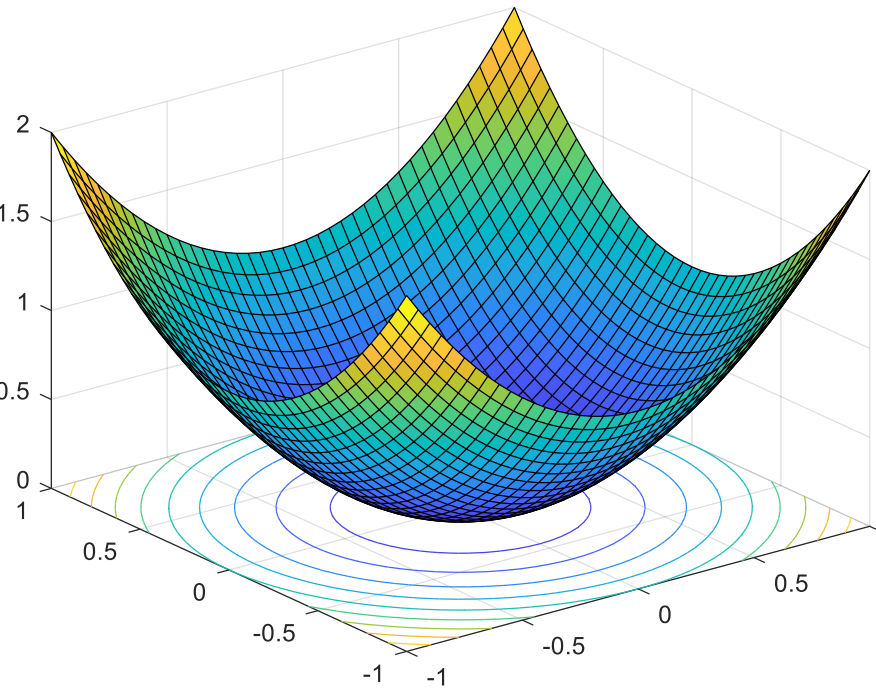
Detailed observations:

- Linear functions are convex
- Any equality constraints must be linear
 - $h(x) = 0 \Leftrightarrow h(x) \geq 0 \text{ AND } h(x) \leq 0$

Convex Programs

minimize $f(x)$

subject to $g_i(x) \leq 0, i = 1, \dots, n,$
where $g_i(x)$ are convex
 $h_j^\top x = 0, j = 1, \dots, m$

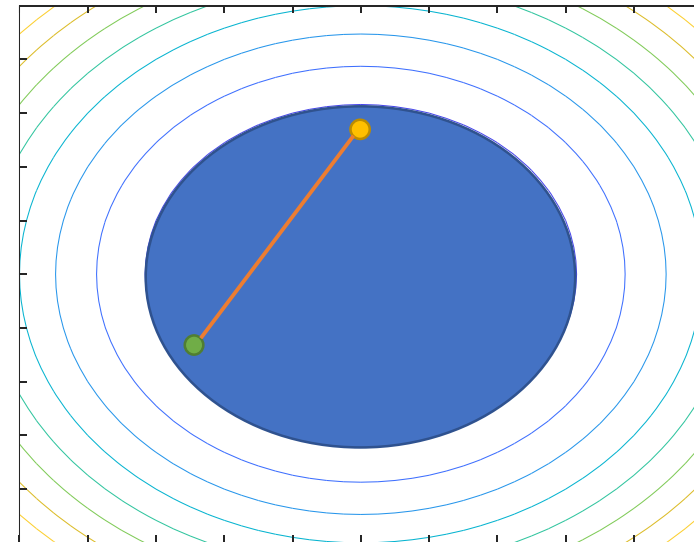
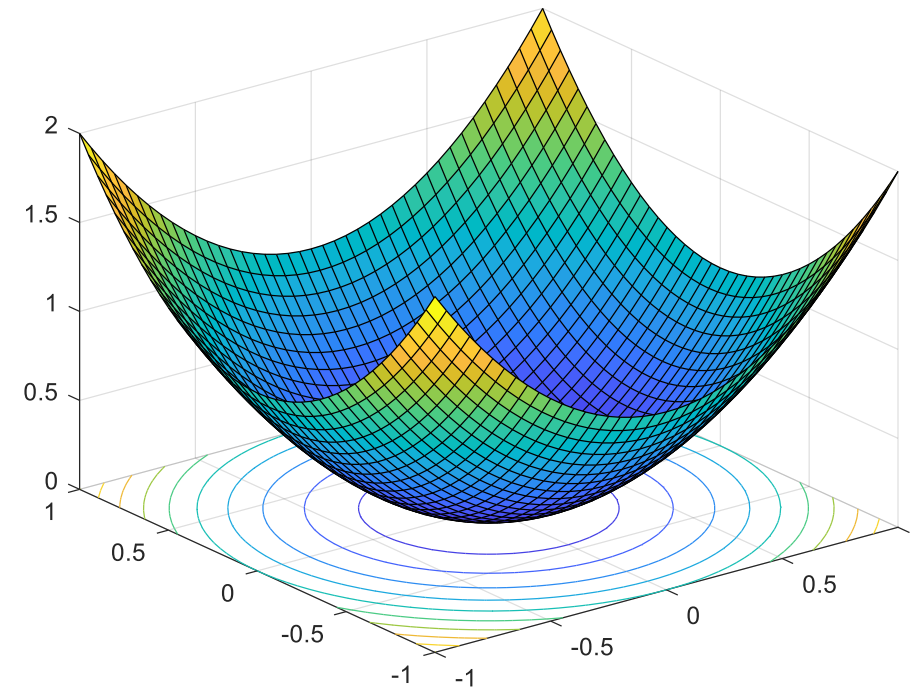


Convex Programs

minimize $f(x)$

subject to $g_i(x) \leq 0, i = 1, \dots, n,$
 where $g_i(x)$ are convex
 $h_j^\top x = 0, j = 1, \dots, m$

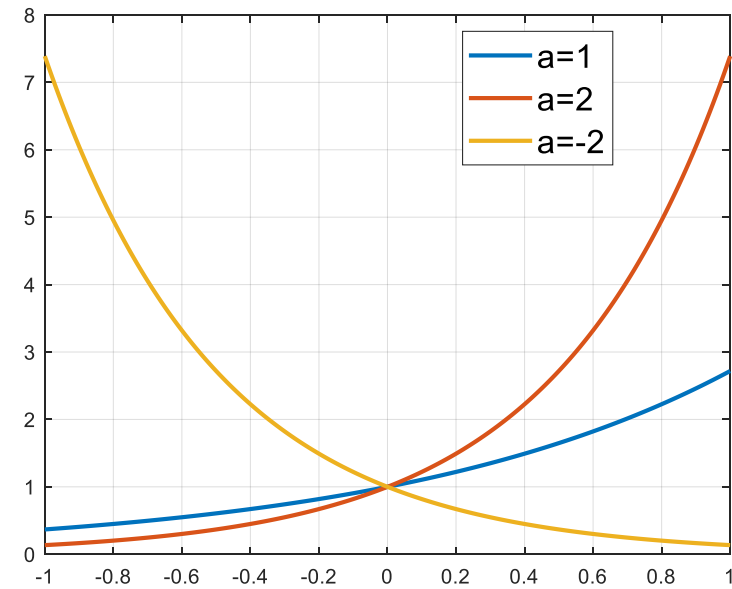
- Local optimum is global!
- Relatively easy to solve using simple algorithms
- When you see an optimization problem, first hope it's convex (although this is almost never true)
 - If an optimization problem is not convex, usually one can only hope for local optimum
- It is useful to recognize convex functions



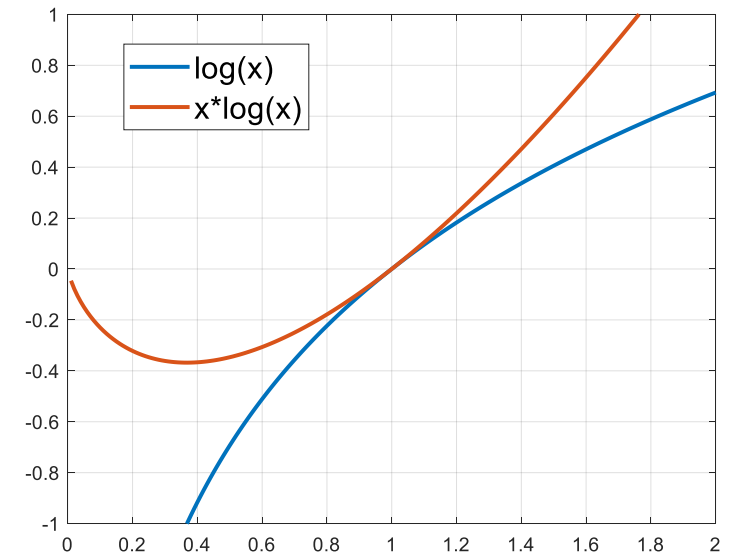
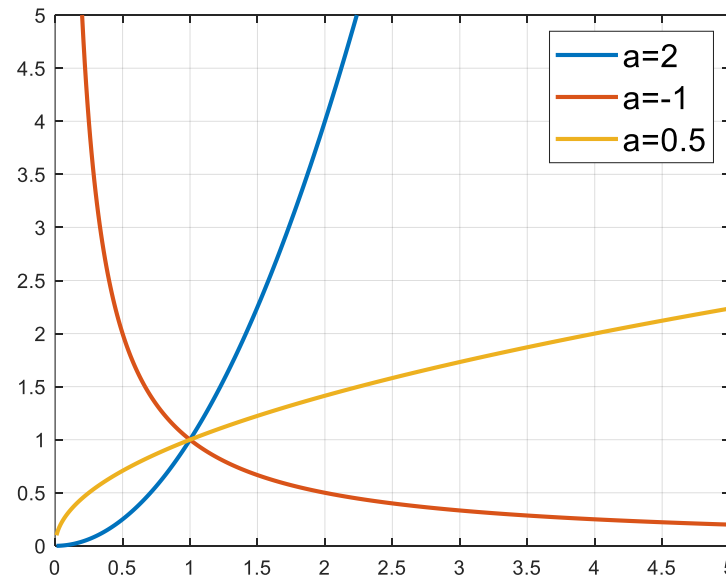
Common Convex Functions on \mathbb{R}

- $f(x) = e^{ax}$ is convex for all $x, a \in \mathbb{R}$
- $f(x) = x^a$ is convex on $x > 0$ if $a \geq 1$ or $a \leq 0$; concave if $0 < a < 1$
- $f(x) = \log x$ is concave
- $f(x) = x \log x$ is convex for $x > 0$ (or $x \geq 0$ if defined to be 0 when $x = 0$)

$$f(x) = e^{ax}$$



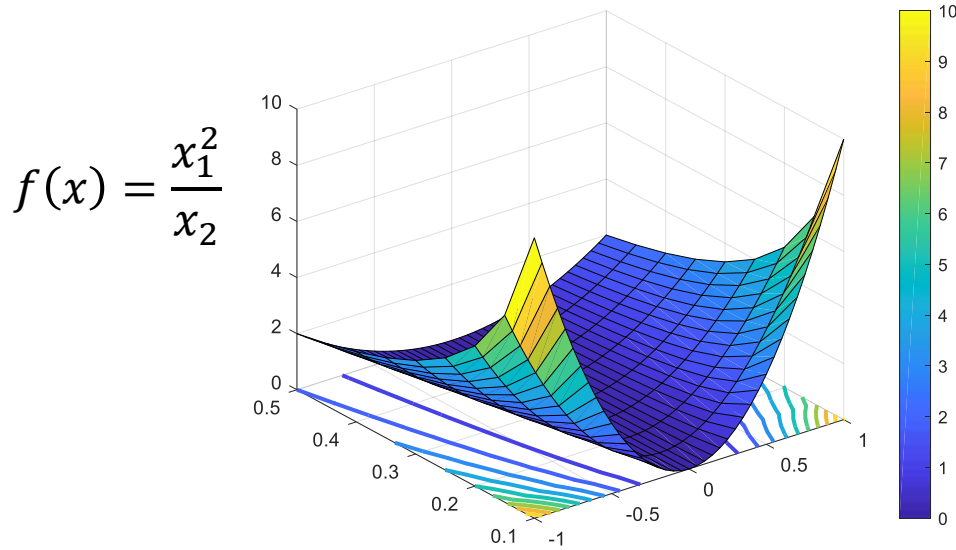
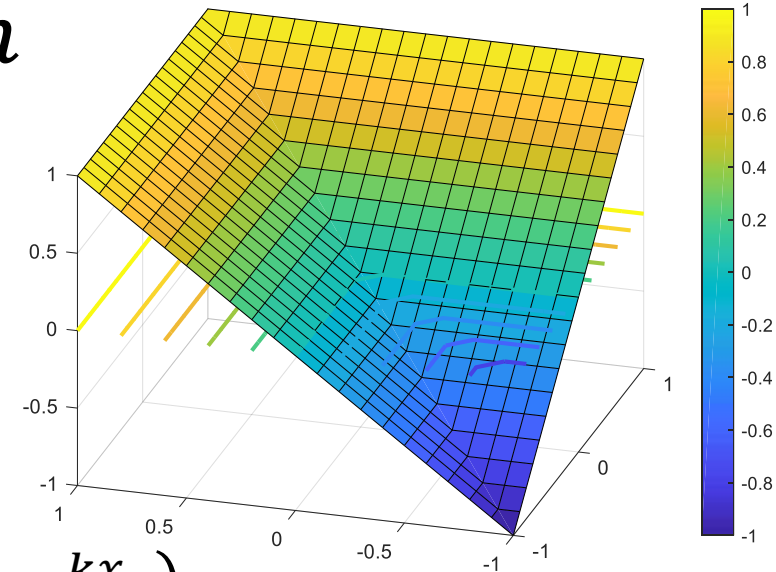
$$f(x) = x^a$$



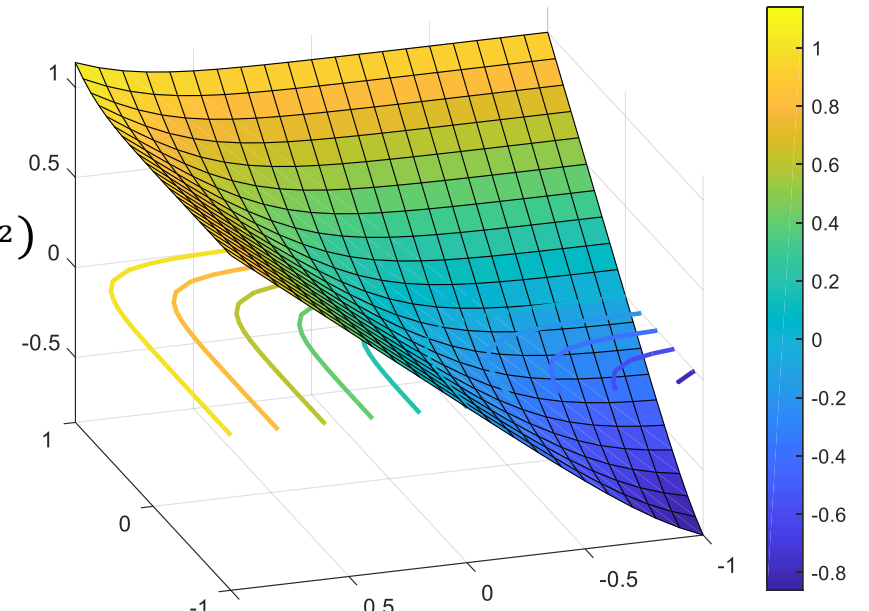
Common Convex Functions on \mathbb{R}^n

- $f(x) = Ax + b$ is convex for any A, b
- Every norm on \mathbb{R}^n is convex
- $f(x) = \max(x_1, x_2, \dots, x_n)$ is convex
- $f(x) = \frac{x_1^2}{x_2}$ (for $x_2 > 0$)
- Log-sum-exp softmax: $f(x) = \frac{1}{k} \log(e^{kx_1} + e^{kx_2} + \dots + e^{kx_n})$
- Geometric mean: $f(x) = (\prod_{i=1}^n x_i)^{\frac{1}{n}}, x_i > 0$

$$f(x_1, x_2) = \max(x_1, x_2)$$



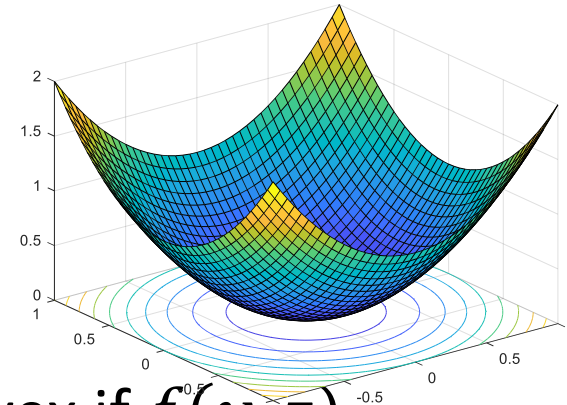
$$f(x) = \frac{1}{5} \log(e^{5x_1} + e^{5x_2})$$



Operations that Preserve Convexity

- Non-negative weighted sum: $\sum_i w_i f_i(x)$ is convex if $f_i(x)$ are convex and $w_i \geq 0$
 - Example: $f(x) = ax^2 + bx^4 + cx^6$, where $a, b, c > 0$
- Composition with affine function: $g(x) = f(Ax + b)$ is convex if $f(x)$ is convex
 - Example: $f(\theta) = \|X\theta - Y\|_2^2$
- Point-wise maximum: $\max(f_1(x), f_2(x))$

Operations that Preserve Convexity



- Point-wise minimum of a function: $g(y) := \min_z f(y, z)$ is convex if $f(y, z)$ is convex (jointly in (y, z))
- Perspective: $g(x, t) := tf\left(\frac{x}{t}\right)$, $t > 0$ is convex if $f(x)$ is convex
 - Example: $\frac{x_1^2}{x_2}$ is convex if $x_2 > 0$, because $f(x_1) = x_1^2$ is convex
- If $g_i: \mathbb{R}^n \rightarrow \mathbb{R}$ are convex, and $h: \mathbb{R}^k \rightarrow \mathbb{R}$ is convex and non-decreasing in each argument, then $h(g_1(x), g_2(x), \dots, g_k(x))$ is convex
 - Example: $\log(e^{g_1(x)} + e^{g_2(x)} + \dots + e^{g_k(x)})$ is convex if g_i are convex, since $\log(e^{x_1} + \dots + e^{x_k})$ is convex
 - More similar composition rules in Boyd and Vandenberghe.

How to check if a function is convex

- Use definition: $f(\theta x + (1 - \theta)y) \leq \theta f(x) + (1 - \theta)f(y)$
- Show $f(y) \geq f(x) + \nabla f(x) \cdot (y - x)$ for differentiable functions
- Show $\nabla^2 f(x) \succcurlyeq 0$ for twice differentiable functions
- **Show f is obtained from simple convex functions and operations that preserve convexity**

Example 1:

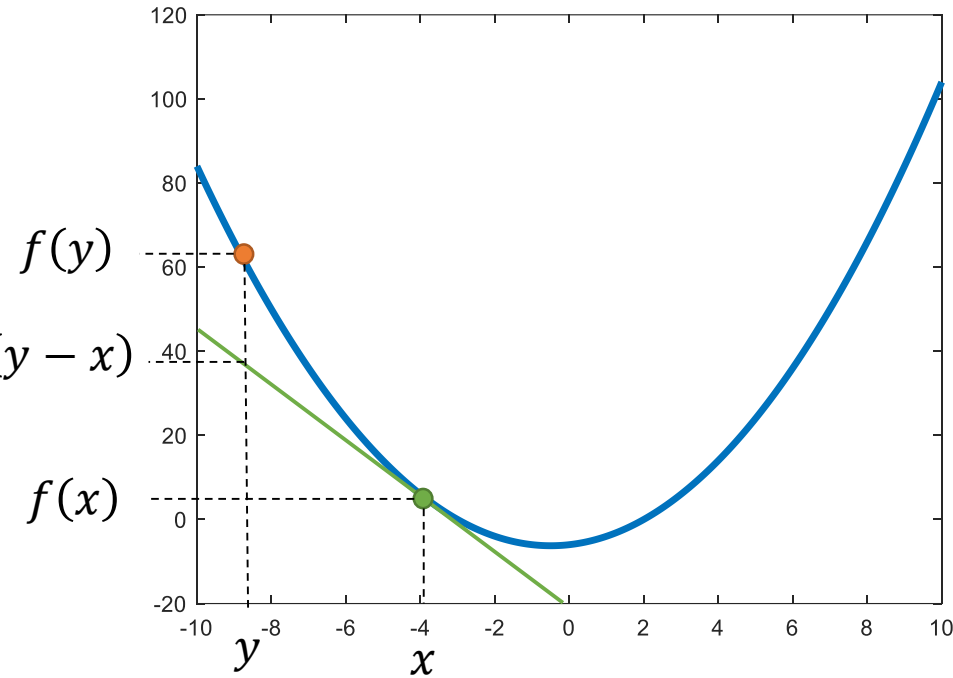
- $f(x) = Ax + b, x \in \mathbb{R}^n$

$$\begin{aligned} f(\theta x + (1 - \theta)y) &= A(\theta x + (1 - \theta)y) + b \\ &= \theta Ax + (1 - \theta)Ay + b \\ &= \theta Ax + (1 - \theta)Ay + \theta b + (1 - \theta)b \\ &= \theta f(x) + (1 - \theta)f(y) \end{aligned}$$

- Equality!
- This means f is also concave (i.e. $-f$ is convex)
- Linear functions are both convex and concave

Example 2:

- $f(x) = x^2 + x - 6$
- Method 1: show $f(y) \geq f(x) + \nabla f(x) \cdot (y - x)$
 - $\nabla f(x) = f'(x) = 2x + 1$



$$\begin{aligned} f(y) - f(x) + f'(x)(y - x) &= y^2 + y - 6 - [x^2 + x - 6 + (2x + 1)(y - x)] \\ &= y^2 + y - [x^2 + x + 2xy - 2x^2 + y - x] \\ &= y^2 + y - [-x^2 + 2xy + y] \\ &= y^2 + x^2 - 2xy \\ &= (x - y)^2 \geq 0 \end{aligned}$$

- Method 2: show $\nabla^2 f(x) \geq 0$

$$\nabla^2 f(x) = f''(x) = 2 \geq 0$$

Example 3:

- $f(x) = \|Ax + b\|_2 + \lambda\|x\|_1$, A is a constant matrix, b is a constant vector, and $\lambda \geq 0$ is a constant scalar.
 - $\|x\|_1$ and $\|x\|_2$ are convex since all norms are convex
 - So, $\|Ax + b\|_2$ is convex, by the rule of affine composition
 - $g(x) = f(Ax + b)$ is convex if $f(x)$ is convex
 - Finally, $\|Ax + b\|_2 + \lambda\|x\|_1$ is convex, by the rule of non-negative weighted sum
 - $\sum_i w_i f_i(x)$ is convex if $f_i(x)$ are convex and $w_i \geq 0$