Index

1. The problem of optimization
2. Direct optimization
3. Descent methods
   – Newton's method
   – Gradient methods
4. Online optimization
5. Constrained optimization
   – Lagrange's method
   – Projected gradients
6. Regularization
7. Convex optimization and Lagrangian duals

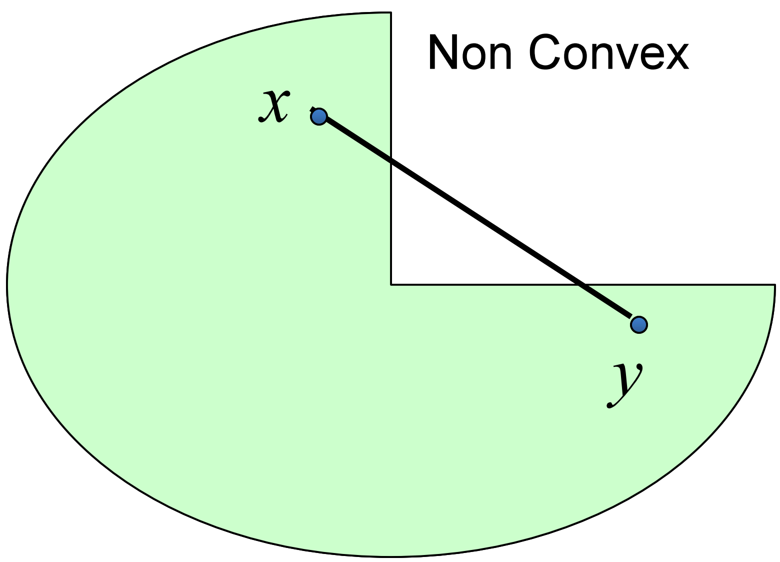
Convex optimization Problems

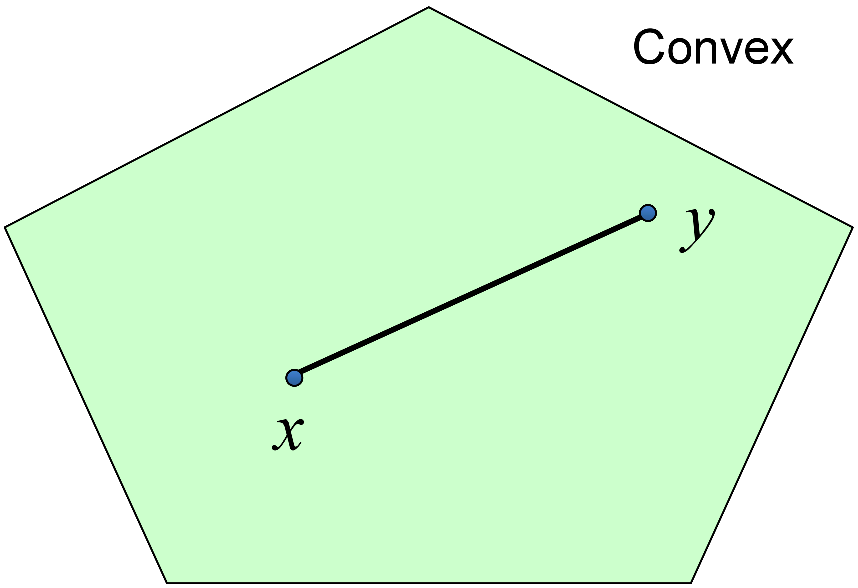
- An convex optimization problem is defined by
  - convex objective function
  - Convex inequality constraints f_i
  - Affine equality constraints h_j

$$
\begin{aligned}
\min_x \quad & f_0(x) \qquad (convex\ function) \\
s.t. \quad & f_i(x) \le 0\ (convex\ sets) \\
& h_j(x) = 0\ (Affine)
\end{aligned}
$$

Convex Sets

- a set $C \in \Re^n$ is convex, if for each $x, y \in C$ and $\alpha \in [0,1]$ then $\alpha x + (1-\alpha) y \in C$

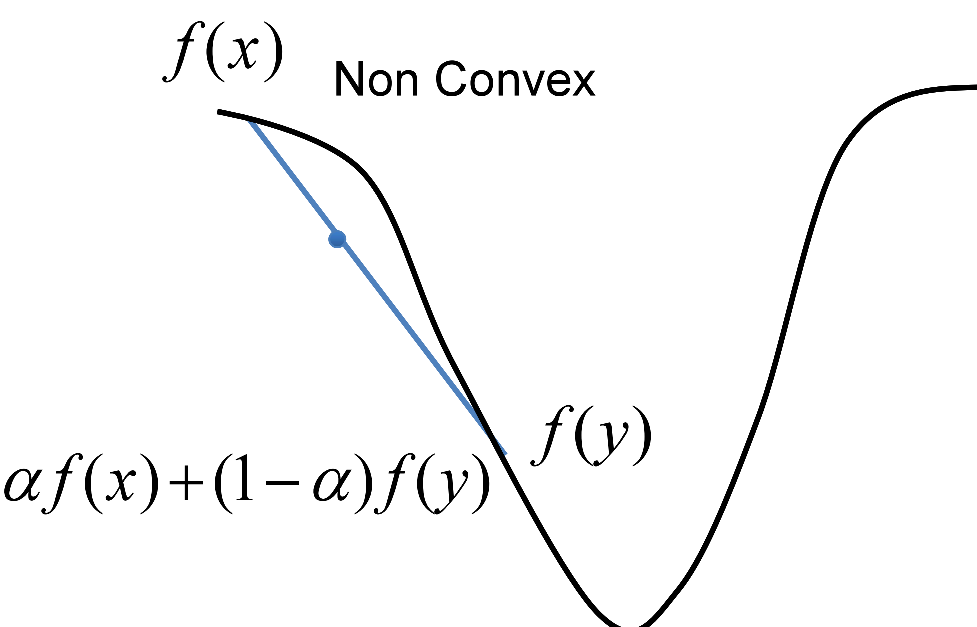


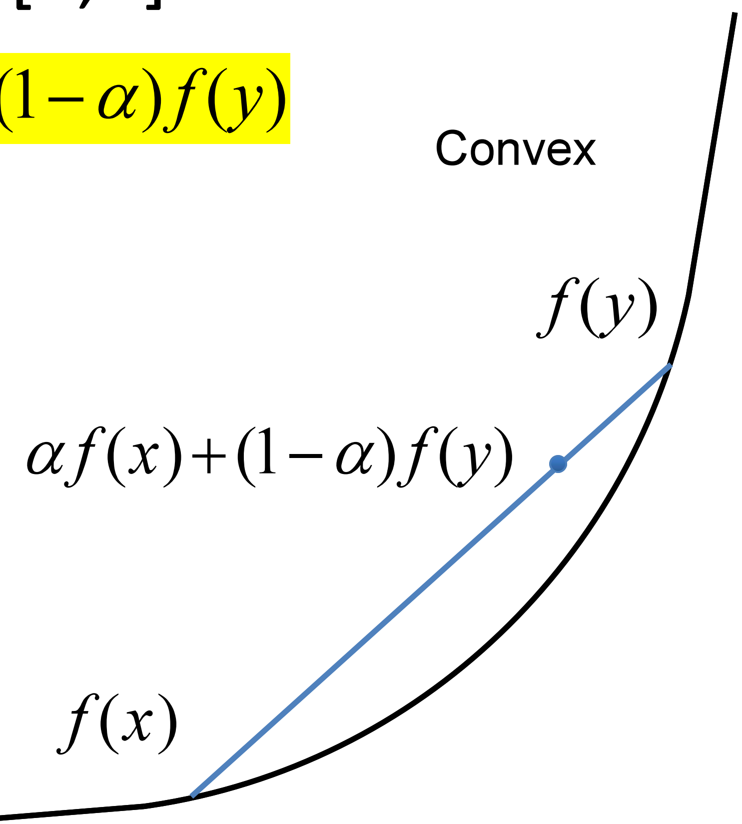


Convex functions

- A function $f:\mathcal{R}^N \longrightarrow \mathcal{R}$ is convex if for each $x, y \in domain(f)$ and $\alpha \in [0,1]$

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



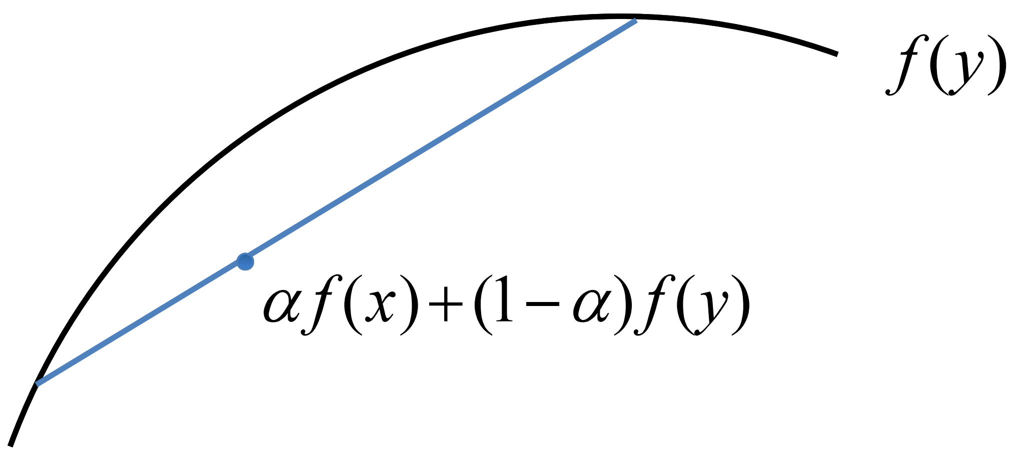


Concave functions

- A function $f: \mathcal{R}^N \longrightarrow \mathcal{R}$ is convex if for each $x, y \in domain(f)$ and $\alpha \in [0,1]$

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

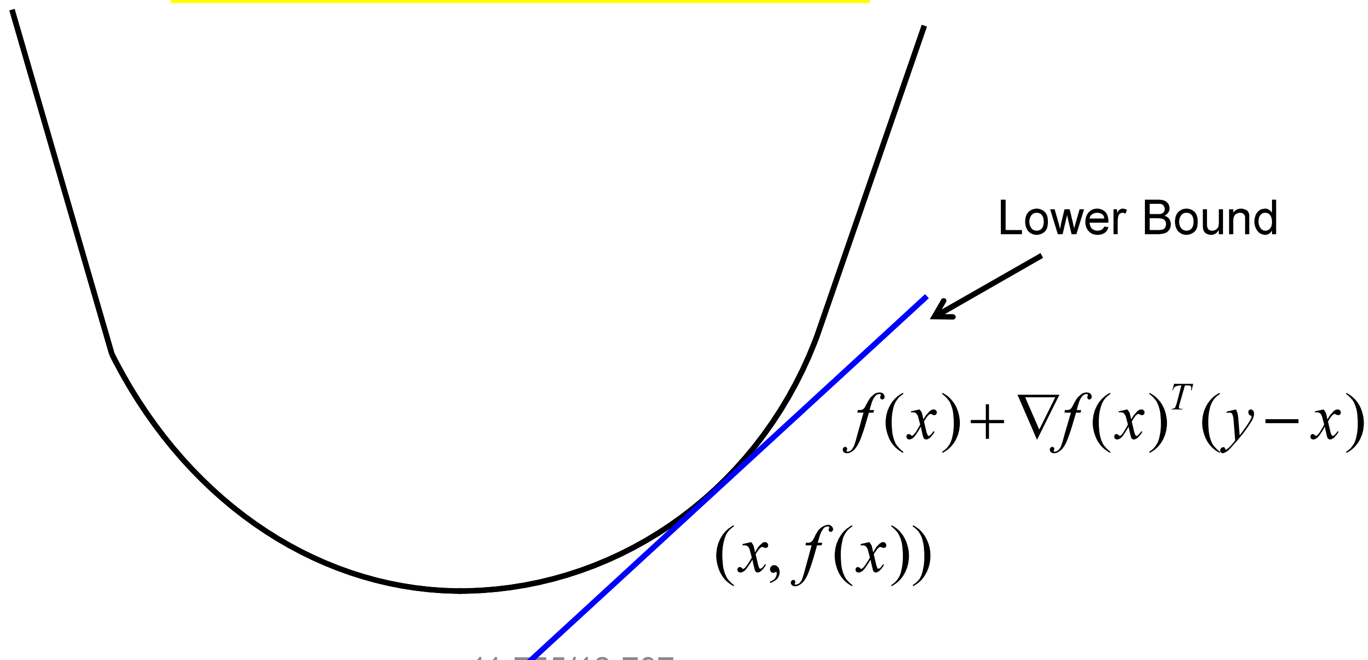
Concave



First order convexity conditions

- A differentiable function $f: \mathcal{R}^N \longrightarrow \mathcal{R}$ is convex if and only if for $x, y \in domain(f)$ the following condition is satisfied

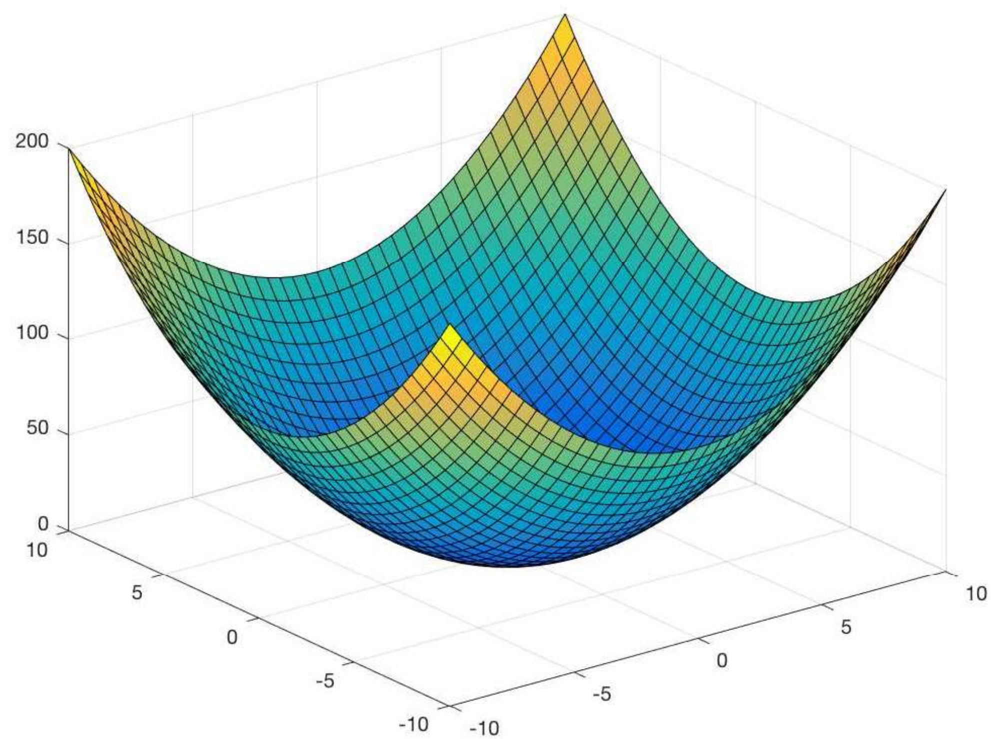
$$f(y) \geq f(x) + \nabla f(x)^t (y - x)$$



Second order convexity conditions

- A twice-differentiable function $f : \mathcal{R}^N \longrightarrow \mathcal{R}$ is convex if and only if for all $x, y \in domain(f)$ the Hessian is superior or equal to zero

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



Properties of Convex Optimization

- For convex objectives over convex feasible sets, the optimum value is unique
  - There are no local minima/maxima that are not also the global minima/maxima

- Any gradient-based solution will find this optimum eventually
  - Primary problem: speed of convergence to this optimum

Lagrange multiplier duality

- Optimization problem with constraints

$$\min_x \; f(x)$$
$$s.t.\; g_i(x) \le 0 \quad i = \{1,\ldots,k\}$$
$$h_j(x) = 0 \quad j = \{1,\ldots,l\}$$

- Lagrange multipliers $\lambda_i \ge 0, \nu \in \Re$

$$L(x,\lambda,\nu) = f(x) + \sum_{i=1}^{k} \lambda_i g_i(x) + \sum_{j=1}^{l} \nu_j h_j(x)$$

- The Dual function

$$\inf_x L(x,\lambda,\nu) = \inf_x \left\{ f(x) + \sum_{i=1}^{k} \lambda_i g_i(x) + \sum_{j=1}^{l} \nu_j h_j(x) \right\}$$

Lagrange multiplier duality

- The Original optimization problem

$$\min_{x}\left\{\sup_{\lambda\geq 0,\nu} L(x,\lambda,\nu)\right\}$$

- The Dual optimization

$$\max_{\lambda\geq 0,\nu}\left\{\inf_{x} L(x,\lambda,\nu)\right\}$$

- Property of the Dual for convex function

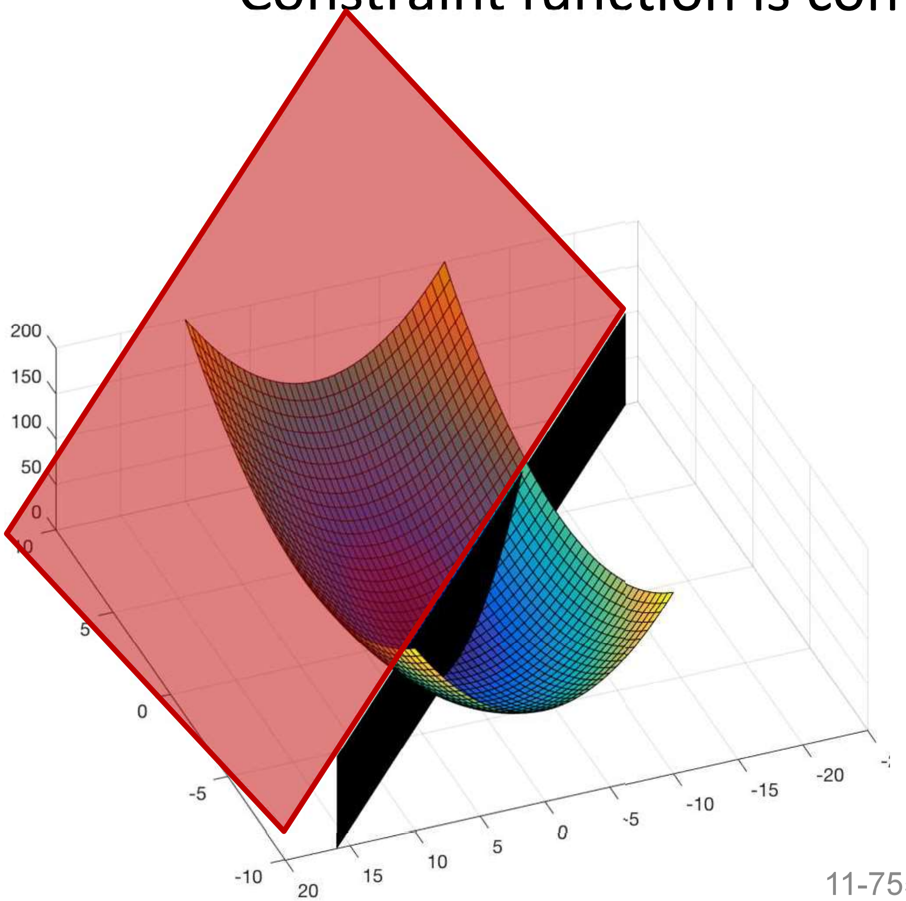
$$\sup_{\lambda\geq 0,\nu}\left\{\inf_{x} L(x,\lambda,\nu)\right\} = f(x^{*})$$

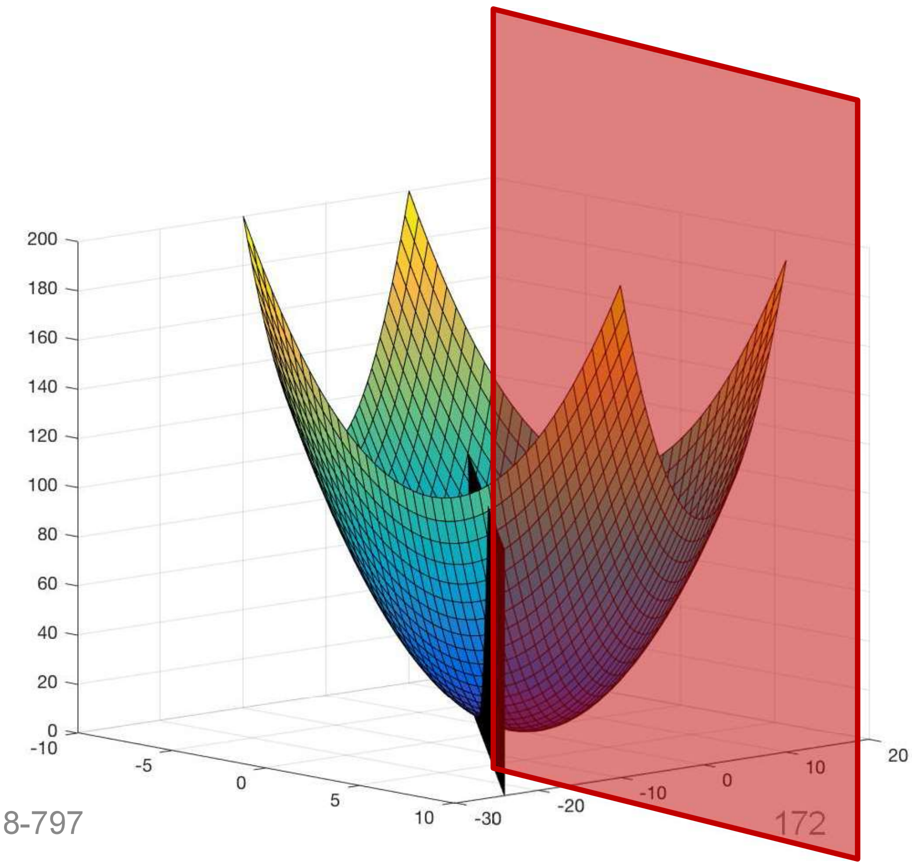
Lagrange multiplier duality

- Previous Example
  - $f(x,y)$is convex
  - Constraint function is convex

$$\min_{x,y} \quad f(x,y) = x^2 + y^2$$

$$s.t. \quad 2x + y \leq -4$$





Lagrange multiplier duality

- Primal system

$$\min_{x,y} \quad f(x,y)=x^2+y^2$$
$$s.t. \quad 2x+y\le -4$$

- Lagrange Multiplier

$$L=x^2+y^2+\lambda(2x+y-4)$$
$$\frac{\partial L}{\partial x}=2x+2\lambda=0\Rightarrow x=-\lambda$$
$$\frac{\partial L}{\partial y}=2y+\lambda=0\Rightarrow y=-\frac{\lambda}{2}$$

- Dual system

$$\max_{\lambda} w(\lambda)=\frac{5}{4}\lambda^2+4\lambda$$
$$s.t. \;\; \lambda\ge 0$$

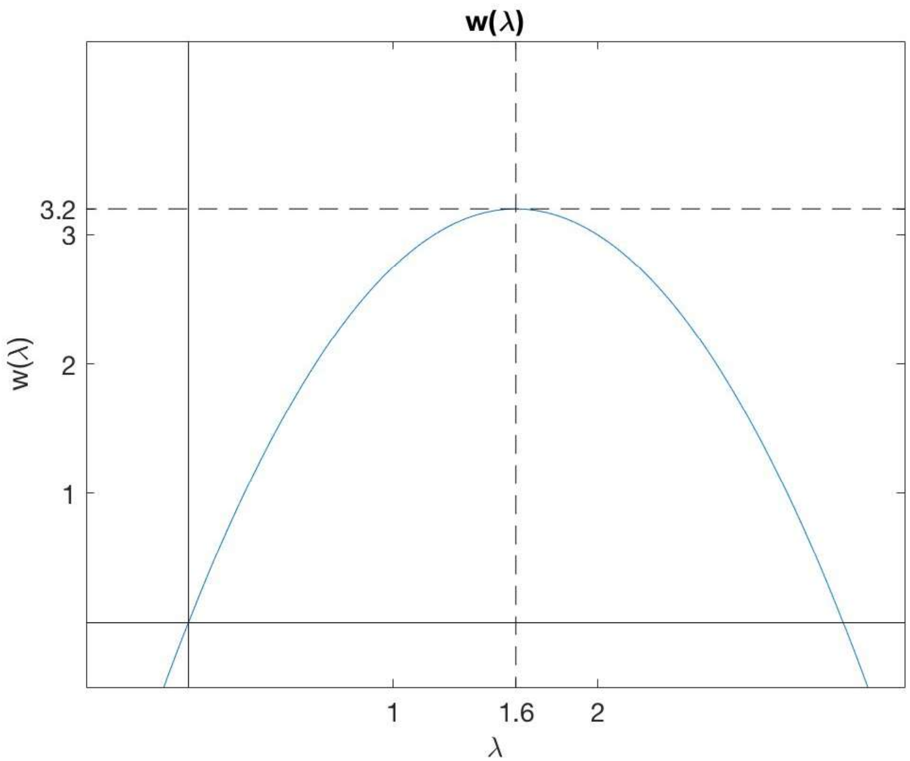
- Property

$$w(\lambda^*)=f(x^*,y^*)$$

Lagrange multiplier duality

- Dual system

$$\max_{\lambda} w(\lambda) = \frac{5}{4}\lambda^2 + 4\lambda$$
$$s.t. \;\; \lambda \geq 0$$



- Concave function
  - Convex function become concave function in dual problem

$$\frac{\partial w}{\partial x} = -\frac{5}{2}\lambda + 4 = 0 \Rightarrow \lambda^* = \frac{8}{5}$$

Lagrange multiplier duality

- Primal system

$$\min_{x,y} \quad f(x,y) = x^2 + y^2$$
$$s.t. \quad 2x + y \le -4$$

- Dual system

$$\max_{\lambda} w(\lambda) = \frac{5}{4}\lambda^2 + 4\lambda$$
$$s.t. \;\; \lambda \ge 0$$

- Evaluating $w(\lambda^*) = f(x^*, y^*)$

$$x^* = -\frac{8}{5}, y^* = -\frac{4}{5}$$
$$f(x^*, y^*) = \left(-\frac{8}{5}\right)^2 + \left(-\frac{4}{5}\right)^2$$
$$f(x^*, y^*) = \frac{16}{5}$$

$$\lambda^* = \frac{8}{5}$$
$$w(\lambda^*) = -\frac{5}{4}\left(\frac{8}{5}\right)^2 + \frac{32}{5}$$
$$w(\lambda^*) = \frac{16}{5}$$