

C477: Mathematical Introduction to Optimisation

Ruth Misener

`r.misener@imperial.ac.uk`

Panos Parpas

`p.parpas@imperial.ac.uk`

Computational Optimisation Group
Department of Computing

Imperial College
London



08 October 2018

Outline

• Topics

- ▶ Global vs local optimisation
- ▶ Neighbourhoods & Other properties of sets
 - ★ Boundary vs interior points; Closed vs open sets; Compact sets
- ▶ Weierstrass Theorem or *How to assert the existence of an optimum*
- ▶ Min vs Argmin
- ▶ Minimisation vs maximisation
- ▶ Defining types of optimisation problems
- ▶ Formulating a nonlinear optimisation problem

• Example

- ▶ Enclosing points

• Reading

- ▶ Chap. 6.1 (Introduction), 19.1 (Introduction), & 4.4 (Neighbourhoods) in *An Introduction to Optimization*, Chong & Zak, Third Edition.

• Acknowledgements

- ▶ Parts of these slides were originally developed by Benoit Chachuat and Panos Parpas. \LaTeX design and proof reading by Miten Mistry. Mistakes by Ruth Misener.

Refresh: Definition: Mathematical Optimisation

Optimisation models (a.k.a. **mathematical programs**) represent problem choices as **decision variables** and seek values that minimise (or maximise) **objective functions** of the decision variables subject to **constraints** on variable values expressing the limits on possible decision choices

$$\begin{array}{ll} \min_{\mathbf{x}} f(\mathbf{x}) & \longleftarrow \text{Objective function} \\ \text{s.t. } \mathbf{h}(\mathbf{x}) = \mathbf{0} & \longleftarrow \text{Equality constraints} \\ \mathbf{g}(\mathbf{x}) \leq \mathbf{0} & \longleftarrow \text{Inequality constraints} \\ \mathbf{x} \in [\mathbf{x}^L, \mathbf{x}^U] = X \subset \mathbb{R}^n & \longleftarrow \text{Variable bounds} \\ \mathbf{h} : \mathbb{R}^n \rightarrow \mathbb{R}^m, \mathbf{g} : \mathbb{R}^n \rightarrow \mathbb{R}^p & \end{array}$$

- **Analytic expressions** of the objective and constraint functions **may or may not be available**
- If we want to find the absolute best set of admissible decisions:
Global optimisation

Defining (Global) Optimality

Feasible Set

The feasible set S (or feasible region) of an optimisation model is the collection of choices for decision variables satisfying **all** of the model

constraints: $S \triangleq \{\mathbf{x} \mid \mathbf{h}(\mathbf{x}) = \mathbf{0}, \mathbf{g}(\mathbf{x}) \leq \mathbf{0}, \mathbf{x}^L \leq \mathbf{x} \leq \mathbf{x}^U\}$

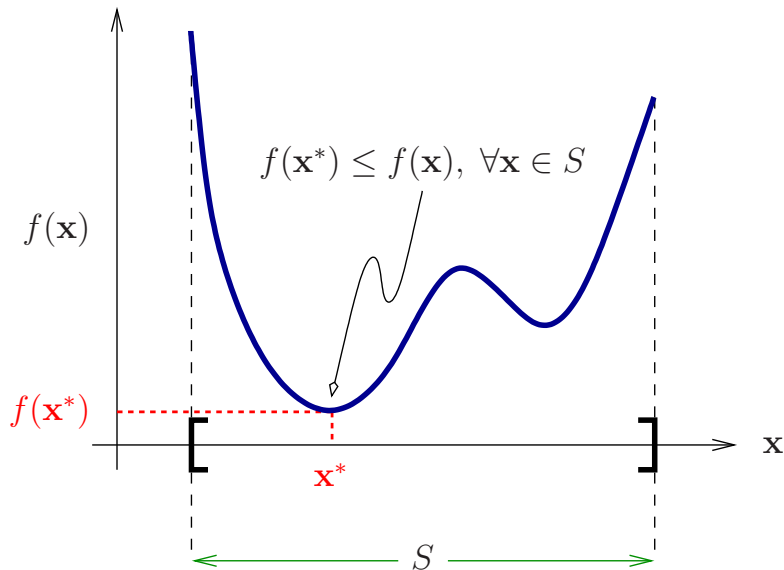
(Global) Optimum

An **optimal solution**, \mathbf{x}^* , is a **feasible** choice for decision variables with objective function value **superior** to any other feasible point. For a

minimisation problem: $f(\mathbf{x}^*) \leq f(\mathbf{x}), \quad \forall \mathbf{x} \in S$

- ❶ The **optimal value** f^* in an optimisation model is the objective function value of any optimal solutions: $f^* = f(\mathbf{x}^*)$ — It is **unique**!
- ❷ But, an optimisation model may have:
 - ▶ a **unique** optimal solution
 - ▶ several **alternative** optimal solutions
 - ▶ **no** optimal solutions (unbounded or infeasible models)

Illustration of a (Global) Minimum, x^*



Defining Local Optimality

Neighbourhood

The **neighbourhood** $N_\delta(\mathbf{x})$ of a point \mathbf{x} consists of all nearby points; that is, all points within a small distance $\delta > 0$ of \mathbf{x} :

$$N_\delta(\mathbf{x}) \triangleq \{\mathbf{y} \mid \|\mathbf{y} - \mathbf{x}\|_2 < \delta\}$$

Local Optimum

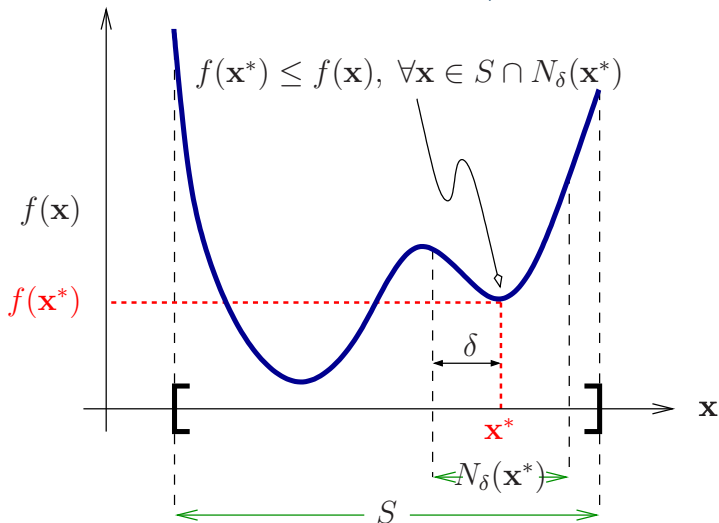
A point \mathbf{x}^* is a **local optimum** for the function $f : \mathbb{R}^n \rightarrow \mathbb{R}$ on the set S if it is feasible ($\mathbf{x}^* \in S$) and if sufficiently small neighbourhoods surrounding it contain no points that are both feasible and superior in objective value:

$$\exists \delta > 0 : \quad f(\mathbf{x}^*) \leq f(\mathbf{x}), \quad \forall \mathbf{x} \in S \cap N_\delta(\mathbf{x}^*)$$

Remarks:

- 1 Global optima are **always** local optima
- 2 Local optima **may not be** global optima

Illustration of a Local Minimum, x^*

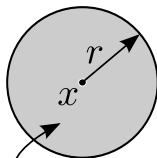
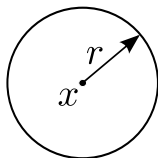


What is a **neighbourhood** $N_\delta(x^*)$? Backing up for definitions ...

A neighbourhood of a point $x \in \mathbb{R}^n$

Neighbourhood

A **neighbourhood** of a point $x \in \mathbb{R}^n$ is the set, $\{y \mid \|x - y\|_2 < r\}$, for some $r > 0$.



Neighbourhood of $x \in \mathbb{R}^n$

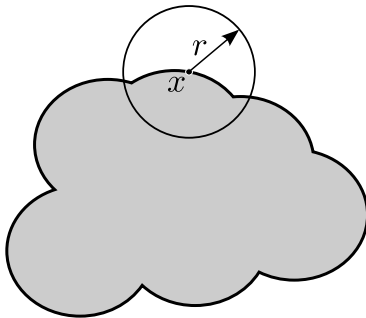
Formal definition

A **neighbourhood** of p in metric space X is a set $N_r(p)$ consisting of all q such that $d(p, q) < r$ for distance metric d and some $r > 0$. The number r is called the **radius** of $N_r(p)$.

Boundary points

Boundary point

A point x is called a **boundary point** if **every** neighbourhood of x contains a point in the set and a point outside the set.



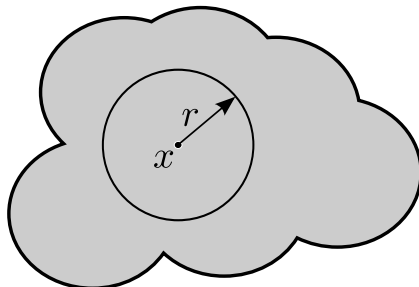
Sanity Check

Give an example of a boundary point?

Interior points

Interior point

A point x is called an **interior point** if all points within **some** neighbourhood of x are contained in the set.



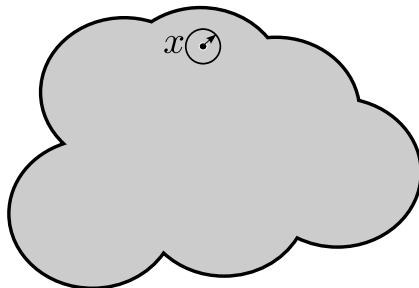
Sanity Check

Give an example of an interior point?

Interior points

Interior point

A point x is called an **interior point** if all points within **some** neighbourhood of x are contained in the set.



Sanity Check

Give an example of an interior point?

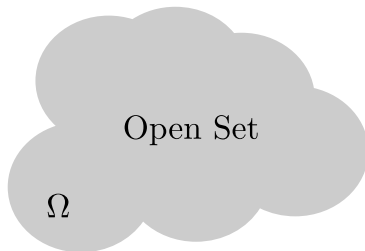
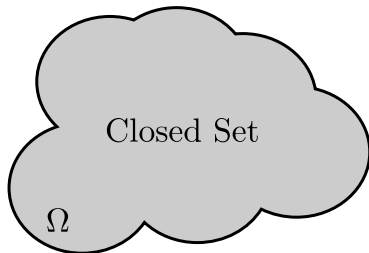
Closed/Open Sets

Closed set

A set is **closed** if it contains its boundary.

Open set

A set is **open** if it contains no boundary points.



Sanity Check

Examples of closed and open sets?

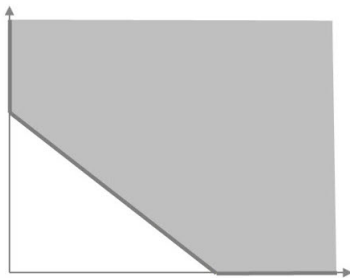
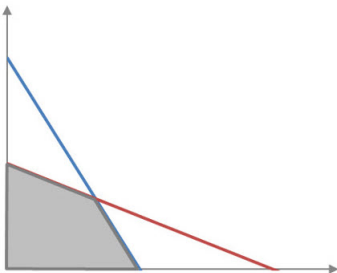
Bounded & Compact Sets

Bounded Set

A set that is contained in a ball of finite radius is **bounded**.

Compact set

A **compact** set is both **closed** and **bounded**.



Sanity Check

Which set is bounded?

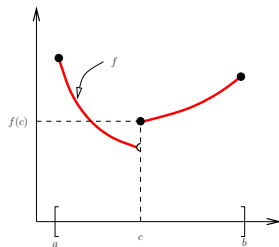
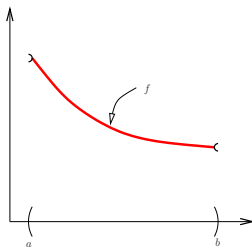
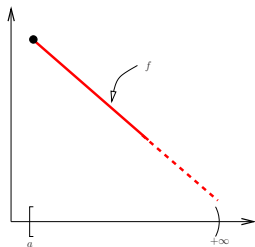
Asserting Existence of Optima

How do we know an optimal solution exists?

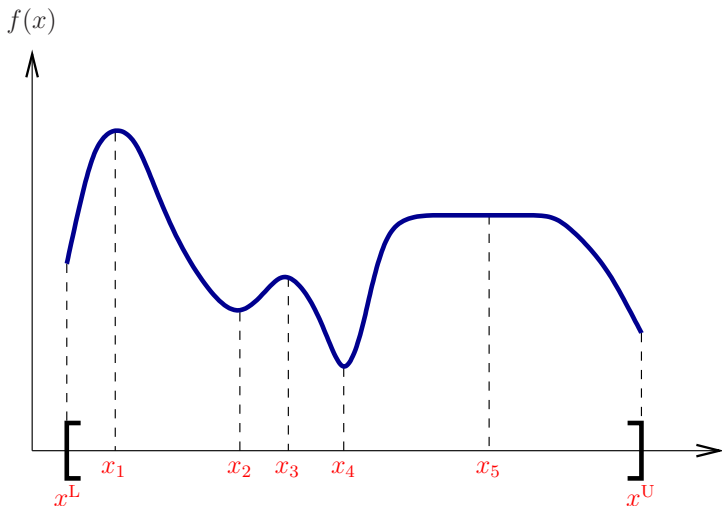
Weierstrass Theorem

Let $S \subset \mathbb{R}^n$ be a nonempty, compact set, and let $f : S \rightarrow \mathbb{R}$ be continuous on S . Then, the problems $\min / \max_{x \in S} f(x)$ attain their optimal values; that is, there exist optimal solution points for either problems.

Why **boundedness** of S ? Why **closedness** of S ? Why **continuity** of f ?



Global vs Local Optima



Sanity Check

Identify the minima & maxima types for f on $S := [x^L, x^U]$

Min vs Argmin

Terminology

If \mathbf{x}^* is a global minimiser we write

$$f^* = \min_{\mathbf{x} \in \Omega} f(\mathbf{x})$$

$$\mathbf{x}^* = \arg \min_{\mathbf{x} \in \Omega} f(\mathbf{x})$$

Examples Assume $x \in \mathbb{R}$. What is $\arg \min f(x)$? What is $\min f(x)$?

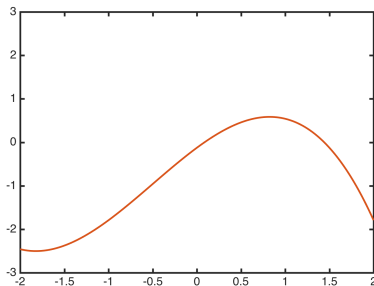
- $f(x) = (x + 1)^2 + 3$
- $f(x) = (x + 1)^2 (x - 1)^2 + 3$

Sanity Check

What is $\arg \min_{\mathbf{x} \in \Omega} f(\mathbf{x})$ on the previous slide.

Minimisation & Maximisation are Equivalent

$$\min_{x \in \Omega} f(x) = -\max_{x \in \Omega} -f(x).$$



$$\min f(x) = -\frac{1}{2}(x+0.5)^2 + 2x - \frac{1}{3}x^3$$

$$x \in [-2, 2]$$

$$\rightarrow x^* \approx -1.823$$

$$f^* \approx -2.502$$

$$-\max -f(x) = \frac{1}{2}(x+0.5)^2 - 2x + \frac{1}{3}x^3$$

$$x \in [-2, 2]$$

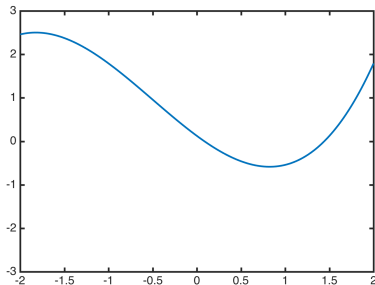
$$\rightarrow x^* \approx -1.823$$

$$f^* \approx -2.502$$

Most optimisation problems in this class will be **minimisation** problems.

Minimisation & Maximisation are Equivalent

$$\min_{x \in \Omega} f(x) = -\max_{x \in \Omega} -f(x).$$



$$\min f(x) = -\frac{1}{2}(x+0.5)^2 + 2x - \frac{1}{3}x^3$$

$$x \in [-2, 2]$$

$$\rightarrow x^* \approx -1.823$$

$$f^* \approx -2.502$$

$$-\max -f(x) = \frac{1}{2}(x+0.5)^2 - 2x + \frac{1}{3}x^3$$

$$x \in [-2, 2]$$

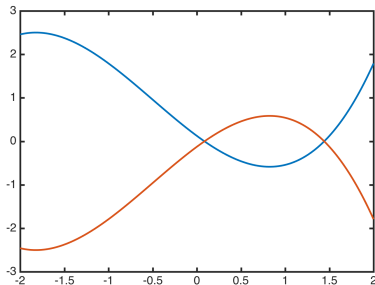
$$\rightarrow x^* \approx -1.823$$

$$f^* \approx -2.502$$

Most optimisation problems in this class will be **minimisation** problems.

Minimisation & Maximisation are Equivalent

$$\min_{x \in \Omega} f(x) = -\max_{x \in \Omega} -f(x).$$



$$\min f(x) = -\frac{1}{2}(x+0.5)^2 + 2x - \frac{1}{3}x^3$$

$$x \in [-2, 2]$$

$$\rightarrow x^* \approx -1.823$$

$$f^* \approx -2.502$$

$$-\max -f(x) = \frac{1}{2}(x+0.5)^2 - 2x + \frac{1}{3}x^3$$

$$x \in [-2, 2]$$

$$\rightarrow x^* \approx -1.823$$

$$f^* \approx -2.502$$

Most optimisation problems in this class will be **minimisation** problems.

Minimisation & Maximisation are Equivalent (proof)

Assume that both the minimum and maximum of f are attained within set Ω and show that:

$$\min_{\mathbf{x} \in \Omega} f(\mathbf{x}) = -\max_{\mathbf{x} \in \Omega} -f(\mathbf{x}).$$

Proof.



Definitions: Types of Optimisation Models (1/4)

Unconstrained Optimisation

$$\min_{\mathbf{x} \in \mathbb{R}^n} f(\mathbf{x})$$

Example Soft-margin Support Vector Machine

$$\min_{\mathbf{w}, b} f(\mathbf{w}, b) = \left[\frac{1}{n} \sum_{i=1}^n \max \left(0, 1 - y_i (\mathbf{w}^\top \mathbf{x}_i + b) \right) \right] + \lambda \|\mathbf{w}\|_2^2$$

Constrained Optimisation

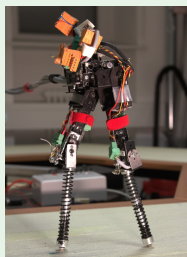
$$\min_{x \in \Omega} f(x)$$

Definitions: Types of Optimisation Models (2/4)

Box-Constrained Optimisation

$$\min_{\mathbf{x} \in [\mathbf{x}^L, \mathbf{x}^U]} f(\mathbf{x})$$

Example Make a robot walk



Objective: optimise performance, e.g., as measured by velocity, within the range of reasonable gait parameters. But there is no closed-form mathematical formulation for the velocity as a function of the input parameters. To solve this problem, the authors use a type of **black-box optimisation** called **Bayesian optimisation**.

<http://wp.doc.ic.ac.uk/sml/project/bayesian-optimization/>

Definitions: Types of Optimisation Models (3/4)

Linear Programming (LP)

$$\begin{aligned}\min \quad & \mathbf{c}^\top \mathbf{x} \\ & \mathbf{A} \mathbf{x} \leq \mathbf{b} \\ & \mathbf{x} \in \mathbb{R}^n\end{aligned}$$

Quadratic Programming (QP)

$$\begin{aligned}\min \quad & \frac{1}{2} \mathbf{x}^\top \mathbf{Q} \mathbf{x} + \mathbf{c}^\top \mathbf{x} \\ & \mathbf{A} \mathbf{x} \leq \mathbf{b} \\ & \mathbf{x} \in \mathbb{R}^n\end{aligned}$$

Portfolio optimisation is often formulated as a QP

Balance expected value of the portfolio's return versus financial risk measures.

Definitions: Types of Optimisation Models (4/4)

Nonlinear Optimisation (NLP)

$$\begin{array}{ll} \min_{\mathbf{x} \in X} f(\mathbf{x}) & \leftarrow \text{Objective function} \\ \text{s.t. } \mathbf{h}(\mathbf{x}) = \mathbf{0} & \leftarrow \text{Equality constraints} \\ \mathbf{g}(\mathbf{x}) \leq \mathbf{0} & \leftarrow \text{Inequality constraints} \\ \mathbf{x} \in [\mathbf{x}^L, \mathbf{x}^U] = X \subset \mathbb{R}^n & \leftarrow \text{Continuous variable bounds} \end{array}$$

$$f : \mathbb{R}^n \rightarrow \mathbb{R}, \mathbf{h} : \mathbb{R}^n \rightarrow \mathbb{R}^m, \mathbf{g} : \mathbb{R}^n \rightarrow \mathbb{R}^p$$

Historical Note

I keep saying *program* when I mean *optimisation problem* because the first well-defined optimisation problems solved *training and logistics schedules*. George Dantzig referred to the solution, or proposed plan, as a *program*.

Mixed-Integer Nonlinear Optimisation (C343 covers mixed-integer optimisation)

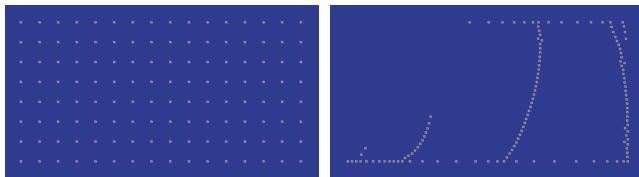
Mixed-Integer Nonlinear Optimisation (MINLP)

$$\begin{aligned} \min_{\mathbf{x} \in X, \mathbf{y} \in Y} \quad & f(\mathbf{x}, \mathbf{y}) && \leftarrow \text{Objective function} \\ \text{s.t.} \quad & \mathbf{h}(\mathbf{x}, \mathbf{y}) = \mathbf{0} && \leftarrow \text{Equality constraints} \\ & \mathbf{g}(\mathbf{x}, \mathbf{y}) \leq \mathbf{0} && \leftarrow \text{Inequality constraints} \\ & \mathbf{x} \in [\mathbf{x}^L, \mathbf{x}^U] = X \subset \mathbb{R}^n && \leftarrow \text{Continuous variable bounds} \\ & \mathbf{y} \in \{0, 1\}^{n_y} && \leftarrow \text{Binary variables} \end{aligned}$$

- Widely used in industry MIP
- Commercial [IBM CPLEX, 2014] MIQP
 - ▶ Passed academic codes from Imperial & CMU in 2015 [ISMP]
- Strong commercial goal MIQCQP
- Wait-&-See MINLP

Other types of optimisation

- Dynamic optimisation is useful for optimisation-based control;
- Partial differential equation-constrained optimisation has loads of engineering applications, e.g. in the *Grantham Institute for Climate Change* at Imperial: <http://arxiv.org/pdf/1304.1768.pdf>
 - ▶ Oceanic tide stream generators could harvest renewable energy;
 - ▶ *Optimisation objective*: How to place the turbines within the site and individually tune them for maximal power output?
 - ▶ *Optimisation constraints*: Need to accurately consider tidal flow, turbine wakes, and the resulting power output.



(a) Initial turbine positions (128 turbines) (b) Optimised turbine positions (128 turbines)

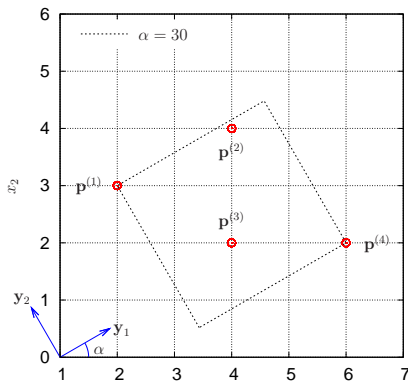
Figure: <http://arxiv.org/pdf/1304.1768.pdf>

A First Example Problem – Statement

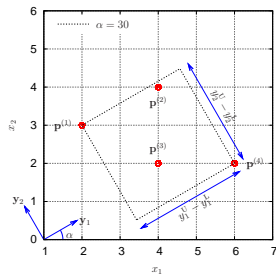
Workshop. Find a rectangle with minimum area enclosing the set of points $\{\mathbf{p}^{(1)} := (2, 3), \mathbf{p}^{(2)} := (4, 4), \mathbf{p}^{(3)} := (4, 2), \mathbf{p}^{(4)} := (6, 2)\}$

➡ Formulate this problem as an NLP

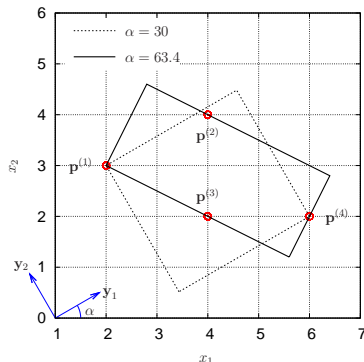
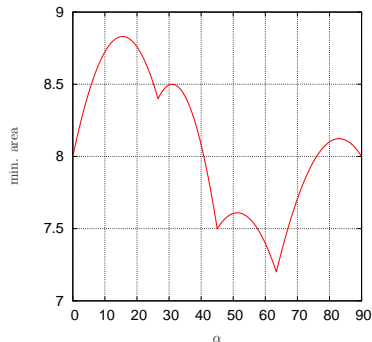
Hint: Consider the coordinate system (y_1, y_2) aligned with the axis of the rectangle, as obtained by rotating (x_1, x_2) by the angle α



A First Example Problem – NLP Formulation



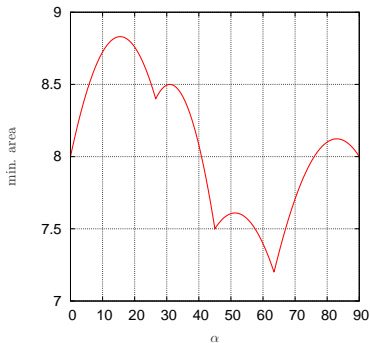
A First Example Problem – Results



● Multi-Start Gradient Descent: 20,000 Random Initial points

Freq	Area	α	y_1^U	y_1^L	y_2^U	y_2^L
32.6%	8.00	0.0	6.000	2.000	4.000	2.000
30.5%	7.20	63.4	5.367	3.578	-0.447	-4.472
15.2%	8.00	90.0	4.000	2.000	-2.000	-6.000
12.1%	7.50	45.0	5.657	3.536	0.707	-2.828
9.7%	8.40	26.6	6.261	3.130	1.789	-0.894

Motivation for next time



What is going on? We are getting different answers depending on where we initialise the algorithm. When is a local optimum also the global solution?

Key concept: **convexity**. If a problem is convex, then a local optimum solution must be a global solution. (More on this Wednesday!!)