# IB Optimisation: Lecture 1

Mike Tehranchi

University of Cambridge

24 April 2020

UNIVERSITY OF CAMBRIDGE

Our typical problem is of the form

$$\text{minimise } f(x) \text{ subject to } g(x) = b, \quad x \in X.$$

- $f : \mathbb{R}^n \to \mathbb{R}$ is the *objective function.*
- $X \subseteq \mathbb{R}^n$ defines a *regional constraint.*
- $g : \mathbb{R}^n \to \mathbb{R}^m$ and $b \in \mathbb{R}^m$ defines $m$ *functional constraints.*

We will use the terminology:

- A *feasible solution* is any $x \in X$ such that $g(x) = b$.
- An *optimal solution* is a feasible solution $x^*$ such that $f(x^*) \leq f(x)$ for all feasible $x$.
- The problem is *feasible* if there exists at least one feasible solution.
- The problem is *bounded* if

$$\inf\{f(x) : \quad g(x) = b, \quad x \in X\} > -\infty$$

We also consider problems of the form

$$\text{P: maximise } f(x) \text{ subject to } g(x) = b, \quad x \in X.$$

- Feasibility of a solution is defined as before
- An *optimal solution* is a feasible solution $x^*$ such that $f(x^*) \geq f(x)$ for all feasible $x$.

This problem is equivalent to

$$\text{P': minimise } -f(x) \text{ subject to } g(x) = b, \quad x \in X.$$

- Problems P and P' have the same set of feasible solutions.
- Problems P and P' have the same set of optimal solutions.

Consider the problem

$$\text{minimise } f(x) \text{ subject to } a \leq x \leq b$$

in the case where $f : \mathbb{R} \to \mathbb{R}$ is twice differentiable..

### Theorem (Necessary conditions for optimality)

*Let $x^*$ be optimal for the problem, and suppose $a < x^* < b$ then $f'(x^*) = 0$*

*Proof.* Let $\varepsilon > 0$ be small enough that both $x^* - \varepsilon$ and $x^* + \varepsilon$ are feasible. Since

$$\frac{f(x^*) - f(x^* - \varepsilon)}{\varepsilon} \leq 0 \leq \frac{f(x^* + \varepsilon) - f(x)}{\varepsilon}$$

Sending $\varepsilon \searrow 0$ yields $f'(x^*) \leq 0 \leq f'(x^*)$ as desired. □

**Theorem (Sufficient conditions for optimality)**

*Suppose that $x^*$ is feasible and $f'(x^*) = 0$. If $f''(x) \geq 0$ for all feasible $x$, then $x^*$ is optimal.*

*Proof.* Let $x$ be a feasible solution. By Taylor's theorem

$$f(x) = f(x^*) + f'(x^*)(x - x^*) + \frac{1}{2} f''(\xi)(x - x^*)^2.$$

where $\xi$ is some point between $x$ and $x^*$. Since $f'(x^*) = 0$ and $f''(\xi) \geq 0$ by assumption, we have $f(x) \geq f(x^*)$. □

## Notational conventions

- a point in $\mathbb{R}^n$ is an $n$-dimensional *column* vector
- If $f$ is differentiable, then $Df(x) \in \mathbb{R}^n$ is the gradient of the function $f$ at the point $x$
- If $f$ is twice differentiable, then $D^2f(x)$ is the $n \times n$ Hessian matrix of second order partial derivatives.

### Definition

A set $X \subseteq \mathbb{R}^n$ is *convex* if for every pair of points $x, y \in X$ and number $0 < p < 1$ we have

$$px + (1 - p)y \in X$$

Let $X \subseteq \mathbb{R}^n$ be convex.

## Definition

A function $f : X \to \mathbb{R}$ is *convex* if for every pair of points $x, y \in X$ and number $0 < p < 1$ we have

$$f(px + (1-p)y) \leq pf(x) + (1-p)f(y)$$

### Theorem (Supporting hyperplane)

*Let $X \subseteq \mathbb{R}^n$ be convex. The function $f : X \to \mathbb{R}$ is convex $\Leftrightarrow$ for every $x \in X$ there exists a vector $\lambda(x) \in \mathbb{R}^n$ such that*

$$f(y) - f(x) \geq \lambda(x)^\top (y - x)$$

*for all $y \in X$.*

*Proof of* $\Leftarrow$. First suppose $\lambda(x)$ exists for all $x$. Fix $y, z \in X$ and $0 < p < 1$, and let $x = py + (1-p)z$. Then

$$f(y) - f(x) \geq \lambda(x)^\top (y - x)$$
$$f(z) - f(x) \geq \lambda(x)^\top (z - x)$$

Hence

$$pf(y) + (1-p)f(z) - f(x) \geq \lambda(x)^\top (py + (1-p)z - x) = 0$$

so $f$ is convex. $\square$

*Proof of* $\Rightarrow$ *when* $f$ *is differentiable*. Now suppose $f$ is convex. By definition, for $0 < p < 1$ we have

$$\frac{f(x + p(y - x)) - f(x)}{p} \leq f(y) - f(x).$$

Now send $p \searrow 0$ and simplify the left-hand side using vector calculus. The vector $\lambda(x) = Df(x)$ satisfies the desired inequality. $\square$

Consider the problem

$$\text{minimise } f(x) \text{ subject to } x \in X$$

where $f : X \to \mathbb{R}$ is differentiable.

### Theorem (Sufficient conditions for optimality)

*Suppose that $x^*$ is feasible and that $Df(x^*) = 0$. If $f$ is convex, then $x^*$ is optimal.*

*Proof.* Let $x$ be feasible. By the supporting hyperplane theorem

$$f(x) - f(x^*) \geq (x - x^*)^\top Df(x^*)$$

But the right-hand side is zero by assumption, hence
$f(x) \geq f(x^*)$. $\square$.

### Definition

A symmetric $n \times n$ matrix is *non-negative definite* if for every $x \in \mathbb{R}^n$ we have $x^\top Ax \geq 0$.

Let $X \subseteq \mathbb{R}^n$ be convex and suppose $f : X \to \mathbb{R}^n$ is twice-differentiable.

### Theorem (Hessian of a convex function)

*If the matrix $D^2 f(x)$ is non-negative definite for all $x$, then the function $f$ is convex.*

*Proof.* For any two $x, y \in X$ we have by Taylor's theorem that

$$f(y) = f(x) + (y - x)^\top Df(x) + \frac{1}{2}(y - x)^\top D^2 f(\xi)(y - x)$$

where $\xi = px + (1 - p)y$ for some $0 < p < 1$. Hence

$$f(y) - f(x) \geq \lambda(x)^\top (y - x)$$

for all $x, y \in X$, where $\lambda(x) = Df(x)$. Then $f$ is convex by the supporting hyperplane theorem. □