Optimisation IB Easter

Main goal:

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

- Objective function $f : \mathbb{R}^n \to \mathbb{R}$
- Regional constraints: $X \subseteq \mathbb{R}^n$
- Functional constraints: $g : \mathbb{R}^n \to \mathbb{R}^m$, $b \in \mathbb{R}^m$

Solutions $x \in \mathbb{R}^n$:

- Feasible solution: satisfy constraints
- Optimal: the feasible solution $x*$ such that $f(x*) \leq f(x)$ for all feasible x (if exists)

1 Convex Optimisation

Convex set: $X \subseteq \mathbb{R}^n$ is convex if for any $x, y \in X$, $t \in [0, 1]$

$$(1-t)x + ty \in X$$

Convex function : f : (convex) $X \to \mathbb{R}$ is convex if for all $t \in [0, 1]$:

$$(1-t)f(x) + tf(y) \geq f((1-t)x + ty)$$

(chord/line segment lie above the function)

Properties:

- Non-negative definite Hessian matrix $D^2 f(x)$ for all $x \implies$ convex f (Taylor's Theorem)
- Supporting hyperplane/graph lies above tangent plane:

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

Conditions for optimum/global min:

- Necessary condition if x^* is also a local min:
  x^* local min $\implies \nabla f(x^*) = 0$

- Sufficient condition:
  $\nabla f(x^*) = 0$ for convex $f \implies x^*$ is global min
  (f is above the horizontal tangent plane)

Strictly convex function: (replace definition of convex with strict inequality)
f : (convex) $X \to \mathbb{R}$ s.t. for all $t \in [0, 1]$:

$$(1-t)f(x) + tf(y) > f((1-t)x + ty)$$

Properties:

- Positive definite Hessian $D^2 f(x)$ for all $x \implies$ strictly convex
- Strictly convex $\implies$ uniqueness of global min (if exists)

Strongly convex: f is m-strongly convex if the function $x \to f(x) - \frac{m}{2}\|x\|^2$ is convex.

For twice differentiable f, f is m-strongly convex
$\iff D^2 f(x) - mI$ is non-negative definite for all x
$\iff z^T D^2 f(x) z \geq m z^T z$ for all z

- Existence of optimal solution:
  If f strongly complex, continuous, then optimal solution exists
  (for large $\|x\|$, $f(x) \geq f(0)$, problem reduces to the compact set, then continuity $\implies f$ bounded and attains optimum)

Computing the optimal solution: algorithms

Gradient descent Newton's method