# Convex and Non-Convex Optimisation

**Aayush Bajaj**

Version 0.1

2025-05-08

# Table of Contents

1. Mathematical Background ........ 3
2. Convexity ........ 4
3. Unconstrained Optimisation ........ 5
   3.1. Standard Form ........ 5
4. Equality Constraints ........ 6
5. Inequality Constraints ........ 7
6. General Constrained Optimisation ........ 8
7. Numerical Methods (unconstrained) ........ 8
8. Penalty Methods ........ 8
9. Optimal Control Theory ........ 9

# 1. Mathematical Background

**Definition 1.1** **Mathspeak**

1. Axiom: A foundational statement accepted without proof. All other results are built ontop.
2. Proposition: A proved statement that is less central than a theorem, but still of interest.
3. Lemma: A helper'' proposition proved to assist in establishing a more important result.
4. Corollary: A statement following from a theorem or proposition, requiring little to no extra proof.
5. Definition: A precise specification of an object, concept or notation.
6. Theorem: A non-trivial mathematical statement proved on the basis of axioms, definitions and earlier results.
7. Remark: An explanatory or clarifying note that is not part of the formal logical chain but gives insight / context.
8. Claim / Conjecture: A statement asserted that requires a proof.

**Definition 1.2** **Vector Norm**

A vector norm on $\mathbb{R}^n$ is a function $\|\cdot\|$ from $\mathbb{R}^n$ to $\mathbb{R}$ such that:

a) $\|\mathbf{x}\| \geq 0, \forall\, \mathbf{x} \in \mathbb{R}^n$ and $\|\mathbf{x}\| = 0 \iff \mathbf{x} = \mathbf{0}$
b) $\|\mathbf{x}+\mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|\ \forall\, \mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ (Triangle Inequality)
c) $\|\alpha\,\mathbf{x}\| = |\alpha|\ \|\mathbf{x}\|\ \forall \alpha \in \mathbb{R}, \mathbf{x} \in \mathbb{R}^n$

**Theorem 1.3** **Cauchy Shwarz-Inequality**

$$|\mathbf{x}^T\,\mathbf{y}| \leq \|\mathbf{x}\|_2\ \|\mathbf{y}\|_2 \tag{1}$$

**Definition 1.4** **Closed and Bounded Sets**

**Definition 1.5** **Functions**

a) Linear:
b) Affine:
c) Quadratic:

**Definition 1.6** **Symmetric**

Definition, plus trace and determinant properties

**Definition 1.7** **Principal Minors**

# 2. Convexity

**Definition 2.1** **Convex Set**

$\Omega \subset \mathbb{R}^n$ is **convex** if $\theta x + (1-\theta)y \in \Omega$ for every $x, y \in \Omega$ and $\theta \in [0, 1]$.

**Proposition 2.2** **Intersection of Convex Sets**

**Definition 2.3** **Extreme Points**

**Definition 2.4** **Convex Combination**

**Definition 2.5** **Convex Hull**

**Theorem 2.6** **Separating Hyperplane**

**Definition 2.7** **Convex Hull**

**Definition 2.8** **Convex / Concave Functions**

A function $f : \Omega \to \mathbb{R}$ (with $\Omega$ convex) is

- **convex** if $f(\theta x + (1-\theta)y) \leq \theta f(x) + (1-\theta)f(y)$;
- **strictly convex** if strict inequality holds whenever $x \neq y$;
- **concave** if $-f$ is convex.

# 3. Unconstrained Optimisation

## 3.1. Standard Form

**Definition 3.1** **Standard Form**

$$\underset{\mathbf{x}\in\Omega}{\text{minimise}}\, f(\mathbf{x}) \tag{2}$$

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

**Theorem 3.2** **First order necessary conditions**

**Definition 3.3** **Stationary point**

**Definition 3.4** **Saddle point**

**Theorem 3.5** **Second order necessary conditions**

**Corollary 3.6** **Local maximiser**

$\bar{\mathbf{x}}$ is a local maximiser $\Longrightarrow \nabla f(\bar{\mathbf{x}}) = \mathbf{0}$ and $\nabla^2 f(\bar{\mathbf{x}})$ negative semi-definite.

Note: As the definiteness of the Hessian changes, so does the nature of the maximiser.

**Theorem 3.7** **Second order sufficient conditions**

# 4. Equality Constraints

**Definition 4.1** **Standard Form**

$$\begin{aligned} \underset{\mathbf{x} \in \Omega}{\text{minimise}} \quad & f(\mathbf{x}) \\ \text{subject to} \quad & \mathbf{c}_i(\mathbf{x}) = 0 \end{aligned} \tag{3}$$

**Definition 4.2** **Lagrangian**

**Definition 4.3** **Regular Point**

**Definition 4.4** **Matrix of Constraint Gradients**

$$A(\mathbf{x}) = [\nabla \mathbf{c}_i(\mathbf{x}) \;\; \dots \;\; \mathbf{c}_m(\mathbf{x})] \tag{4}$$

**Definition 4.5** **Jacobian**

$$\begin{aligned} J(\mathbf{x}) &= A(\mathbf{x})^T \\ &= \begin{bmatrix} \nabla \mathbf{c}_i (\mathbf{x})^T \\ \vdots \\ \mathbf{c}_m (\mathbf{x})^T \end{bmatrix} \end{aligned} \tag{5}$$

**Proposition 4.6** **First order necessary optimality conditions**

**Corollary 4.7** **Constrained Stationary Point**

**Proposition 4.8** **Second order sufficient conditions**

# 5. Inequality Constraints

**Definition 5.1** **Standard Form**

$$
\begin{aligned}
\underset{\mathbf{x}\in\Omega}{\text{minimise}} \quad & f(\mathbf{x}) \\
\text{subject to} \quad & \mathbf{c}_i(\mathbf{x}) = 0, \qquad i = 1, ..., m_E \\
& \mathbf{c}_i(\mathbf{x}) \le 0, \qquad i = m_E + 1, ..., m
\end{aligned}
\tag{6}
$$

**Definition 5.2** **Convex Problem**

The problem (NLP) is a standard form convex optimisation problem if the objective function $f$ is convex on the feasible set, $\mathbf{c}_i$ is affine for each $i \in$ E, and $\mathbf{c}_i$ is convex for each $i \in \mathcal{I}$.

**Definition 5.3** **Active Set**

The set of active constraints at a feasible point $\mathbf{x}$ is $\mathcal{A}(\mathbf{x}) = \{i \in 1, ..., m : \mathbf{c}_i(\mathbf{x}) = 0\}$ **Note** that this concept only applies to inequality constraints.

**Definition 5.4** **Regular Point**

**Proposition 5.5** **Constrained Stationary Point**

**Theorem 5.6** **Karush Kuhn Tucker (KKT) necessary optimality conditions**

kkt generalises lagrange multipliers

**Theorem 5.7** **Second-order sufficient conditions for strict local minimum**

**Theorem 5.8** **KKT sufficient conditions for global minimum**

**Theorem 5.9** **Wolfe Dual Problem**

strong duality, weak duality?

**Note** **Reduced Hessian**

The reduced Hessian $W_Z^*$ is the projection of the Lagrandian's Hessian onto the tangent space of the constraints at the point $x^*$

## 6. General Constrained Optimisation

why does the reduced hessian exist for both? is there any difference when solving?

## 7. Numerical Methods (unconstrained)

**Definition 7.1** **Rates of convergence of iterative methods**

**Algorithm 7.2** **Line Search Algorithms**

**Algorithm 7.3** **Steepest Descent Method**

**Algorithm 7.4** **Newton's Method**

**Algorithm 7.5** **Conjugate Gradient Method**

## 8. Penalty Methods

**Definition 8.1** **Penalty function**

**Remark**

a) $c : \mathbb{R}^n \to \mathbb{R}$ is a convex function $\Longrightarrow \max \{\mathbf{c}(\mathbf{x}), 0\}^2$ is a convex function

b) $\frac{\partial}{\partial x_i}[\max\{\mathbf{c}(\mathbf{x}), 0\}]^2 = 2 \max\{\mathbf{c}(\mathbf{x}), 0\} \frac{\partial}{\partial x_i}$

**Theorem 8.2** **Convergence Theorem**

# 9. Optimal Control Theory

**Definition 9.1** **Standard Form**

**Definition 9.2** **Hamiltonian**

break this up so it is understandable

**Definition 9.3** **Co-state Equations**

**Theorem 9.4** **Pontryagin Maximum Principle**

partially free target

non-autonmous problem