Optimization
MATH3161 UNSW

Jeremy Le

2024T1

Contents

1 Pre-Requisite Knowledge **2**

2 Optimization - What is it? **2**

2.1 Mathematics of Optimization 2
2.2 Optima and Optimizers 3
2.3 Calculus Aspects 4
2.4 Matrices 5

1 Pre-Requisite Knowledge

Single-Variable Calculus Solution Methods for First and Second-order Ordinary (Linear) Differential Equations

Several Variable Calculus Partial Differentiation, Gradients, Taylor's Series, Normal and Tangent Lines.

Matrix Algebra Vectors, Linear Equations and Matrices, Inverses and Determinants, Subspaces and Rank, Eigenvalues and Eigenvectors, Solving Systems of Ordinary Differential Equations, Quadratic Forms.

2 Optimization - What is it?

Optimization/Optimisation Optimization is a process that finds the "best" possible solutions from a set of feasible solutions. When you optimize something, you are "making it best".

What is an optimization problem? An optimization problem is a mathematical problem of finding the best possible solution from a set of feasible solutions. It has the form of minimizing (or maximizing) an objective function subject to constraints.

2.1 Mathematics of Optimization

Outline

- Mathematical model - Model formulation
- Characterising optima $\Longleftrightarrow$ Optimality principles
- Finding optima $\Rightarrow$ Numerical methods
- Convexity
- Duality
- Maximum principle

Decision variables: what can you change

- Finite dimensional $\mathbf{x} \in \mathbb{R}^n$, Number of variables n
- Infinite dimensional: the control

Objective A mathematical function of the variables quantifying the idea of "best".

- Finite dimensional: variables $\mathbf{x} \in \mathbb{R}^n$, OBjective function $f : \mathbb{R}^n \to \mathbb{R}$
- Infinite dimensional: $f : C([0,T] \to \mathbb{R})$, variables $u \in C([0,T])$
- Co-domain of objective function must be ordered (total order)

Constraints Describe restrictions on the allowable values of variables. Constraint structure for variables $\mathbf{x} \in \mathbb{R}^n$.

- Equality constraints
- Inequality constraints
- Feasible region
- Unconstrained
- Standard formulation
- Algebraic structure of constraints

Standard formulation The standard formulation of a continuous finite dimensional optimization is

$$\begin{aligned} \text{Minimize} \quad & f(\mathbf{x}) \\ \mathbf{x} \in \mathbb{R}^n & \\ \text{subject to} \quad & c_i(\mathbf{x}) = 0, \quad i = 1, \dots, m_E; \\ & c_i(\mathbf{x}) \leq 0, \quad i = m_E + 1, \dots, m \end{aligned}$$

Vector Norm A vector norm of $\mathbb{R}^n$ is a function $\|.\|$ from $\mathbb{R}^n$ to $\mathbb{R}$ such that

1. $\|\mathbf{x}\| \geq 0$ for all $\mathbf{x} \in \mathbb{R}^n$ and $\|\mathbf{x}\| = 0 \iff \mathbf{x} = \mathbf{0}$.
2. $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$ for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$. (Triangle Inequality)
3. $\|\alpha \mathbf{x}\| = |\alpha| \, \|\mathbf{x}\|$ for all $\alpha \in \mathbb{R}, \mathbf{x} \in \mathbb{R}^n$.

2.2 Optima and Optimizers

Global minimum and maximizer A point $\mathbf{x}* \in \Omega$ is a global minimizer or maximizer of $f(\mathbf{x})$ over $\Omega \subseteq \mathbb{R}^n \iff f(\mathbf{x}^*) \leq$ or $\geq f(\mathbf{x})$ for all $\mathbf{x} \in \Omega$. The global minimum is $f(\mathbf{x}^*)$.

Strict global minimum and maximizer A point $\mathbf{x}* \in \Omega$ is a strict global minimizer or maximizer of $f(\mathbf{x})$ over $\Omega \subseteq \mathbb{R}^n \iff f(\mathbf{x}^*) <$ or $> f(\mathbf{x})$ for all $\mathbf{x} \in \Omega, \mathbf{x} \neq \mathbf{x}^*$.

Local minimum and maximizer A point $\mathbf{x}* \in \Omega$ is a local minimizer or maximizer of $f(\mathbf{x})$ over $\Omega \subseteq \mathbb{R}^n \iff$ there exists a $\delta > 0$ such that $f(\mathbf{x}^*) \leq$ or $\geq f(\mathbf{x})$ for all $\mathbf{x} \in \Omega$ with $\|\mathbf{x} - \mathbf{x}^*\| \leq \delta$. Then $f(\mathbf{x}^*)$ is a local minimum.

Strict local minimum and maximizer A point $\mathbf{x}* \in \Omega$ is a strict local minimizer or maximizer of $f(\mathbf{x})$ over $\Omega \subseteq \mathbb{R}^n \iff$ there exists a $\delta > 0$ such that $f(\mathbf{x}^*) <$ or $> f(\mathbf{x})$ for all $\mathbf{x} \in \Omega$ with $0 < \|\mathbf{x} - \mathbf{x}^*\| \leq \delta$.

Extrema The global/local extreme of f over Ω are all the global/local minima and all the global/local maxima.

Existence of a global extrema Let Ω be a compact set and let f be continuous on Ω. Then the global extrema of f over Ω exist.

Relaxation If $f : \mathbb{R}^n \to \mathbb{R}$ and $\bar{\Omega} \subseteq \Omega$ then

$$\min_{\mathbf{x} \in \Omega} f(\mathbf{x}) \leq \min_{\mathbf{x} \in \bar{\Omega}} f(\mathbf{x})$$

Thus, the minimum value of the relaxation problem $\leq$ the minimum value of the original problem.

2.3 Calculus Aspects

Graident Let $f : \mathbb{R}^n \to \mathbb{R}$ be continuously differentiable. The graident $\boldsymbol{\nabla} f : \mathbb{R}^n \to \mathbb{R}^n$ of f at $\mathbf{x}$ is

$$\boldsymbol{\nabla} f(\mathbf{x}) = \begin{bmatrix} \frac{\partial f(\mathbf{x})}{\partial x_1} \\ \frac{\partial f(\mathbf{x})}{\partial x_2} \\ \vdots \\ \frac{\partial f(\mathbf{x})}{\partial x_n} \end{bmatrix}$$

Hessian Let $f : \mathbb{R}^n \to \mathbb{R}$ be continuously differentiable. The Hessian $\boldsymbol{\nabla}^2 f : \mathbb{R}^n \to \mathbb{R}^{n \times n}$ of f at $\mathbf{x}$ is

$$\begin{bmatrix} \dfrac{\partial^2 f(\mathbf{x})}{\partial x_1^2} & \dfrac{\partial^2 f(\mathbf{x})}{\partial x_1 \partial x_2} & \cdots & \dfrac{\partial^2 f(\mathbf{x})}{\partial x_1 \partial x_n} \\ \dfrac{\partial^2 f(\mathbf{x})}{\partial x_2 \partial x_1} & \dfrac{\partial^2 f(\mathbf{x})}{\partial x_2^2} & \cdots & \dfrac{\partial^2 f(\mathbf{x})}{\partial x_n 2 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \dfrac{\partial^2 f(\mathbf{x})}{\partial x_n \partial x_1} & \dfrac{\partial^2 f(\mathbf{x})}{\partial x_n \partial x_2} & \cdots & \dfrac{\partial^2 f(\mathbf{x})}{\partial x_n^2} \end{bmatrix}$$

Linear and Quadratic Functions Let $f_0 \in \mathbb{R}, \mathbf{g} \in \mathbb{R}^n$ and $G \in \mathbb{R}^{n \times n}, G$ symmetric, be fixed. Find the gradient $\boldsymbol{\nabla} f(\mathbf{x})$ and Hessian $\boldsymbol{\nabla}^2 f(\mathbf{x})$ for the

- Linear function $f(\mathbf{x}) = \mathbf{g}^T b\mathbf{x}$; Affine function $f(\mathbf{x}) = \mathbf{g}^x \mathbf{x} + f_0$
- Quadratic function $f(\mathbf{x}) = \frac{1}{2}\mathbf{x}^T G\mathbf{x} + \mathbf{g}^T\mathbf{x} + f_0$

2.4 Matrices

Positive definite matrcies A real square matrix $A \in \mathbb{R}^{n \times n}$ is

- positive definite $\iff$ $\mathbf{x}^T A\mathbf{x} > 0$ for all $\mathbf{x} \in \mathbb{R}^n, \mathbf{x} \neq 0$
- positive semi-definite $\iff$ $\mathbf{x}^T A\mathbf{x} \geq 0$ for all $\mathbf{x} \in \mathbb{R}^n$
- negative definite $\iff$ $\mathbf{x}^T A\mathbf{x} < 0$ for all $\mathbf{x} \in \mathbb{R}^n, \mathbf{x} \neq 0$
- negative semi-definite $\iff$ $\mathbf{x}^T A\mathbf{x} \leq 0$ for all $\mathbf{x} \in \mathbb{R}^n$
- indefinite $\iff$ there exits $\mathbf{x}_0, \mathbf{y}_0 \in \mathbb{R}^n : \mathbf{x}_0^T A\mathbf{x}_0 > 0$ and $\mathbf{y}_0^T A\mathbf{y}_0 < 0$