Numerical Optimisation and Large–Scale Systems MA40050

Pranav Singh
Academic Year 2019/20 – Semester 2

Department of Mathematical Sciences
University of Bath

What is nonlinear programming?

Nonlinear (constrained) optimisation $\equiv$ nonlinear programming

$$\min_{\mathbf{x}\in\mathbb{R}^N} f(\mathbf{x}) \quad \text{subject to} \quad \mathbf{c}_{\mathcal{E}}(\mathbf{x}) = 0 \quad \text{and} \quad \mathbf{c}_{\mathcal{I}}(\mathbf{x}) \geq 0$$

- objective function $f : \mathbb{R}^N \longrightarrow \mathbb{R}$
- constraints $\mathbf{c}_{\mathcal{E}} : \mathbb{R}^N \longrightarrow \mathbb{R}^{M_e}$ $(M_e \leq N)$ and

  $\mathbf{c}_{\mathcal{I}} : \mathbb{R}^N \longrightarrow \mathbb{R}^{M_i}$

An Example

Optimisation of
a high-pressure
gas network

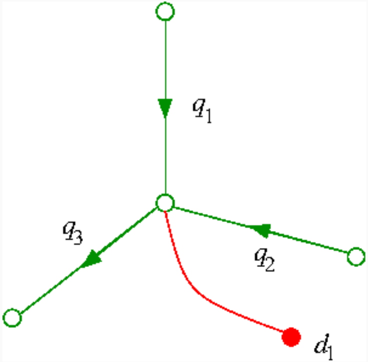
British Gas (Transco)
Oxford University
RAL

Transco
National
Transmission
System

Node Equations

$$q_1 + q_2 - q_3 - d_1 = 0$$

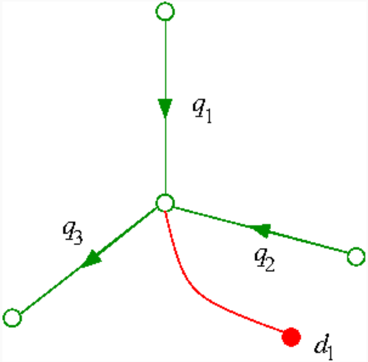
where q_i flows
d_i demands



Node Equations

$$q_1 + q_2 - q_3 - d_1 = 0$$

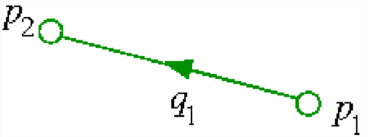
where q_i flows
d_i demands



In general: $A\mathbf{q} - \mathbf{d} = \mathbf{0}$

- linear
- sparse
- structured

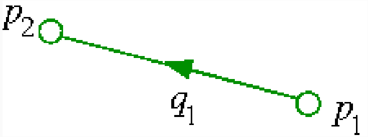
Pipe Equations



$$p_2^2 - p_1^2 + k_1 q_1^{2.8359} = 0$$

where p_i pressures
q_i flows
k_i constants

Pipe Equations



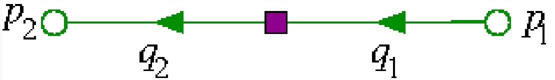
$$p_2^2 - p_1^2 + k_1 q_1^{2.8359} = 0$$

where p_i pressures
q_i flows
k_i constants

In general: $A(\mathbf{p}) + \mathrm{diag}(k_i q_i) = \mathbf{0}$

- non-linear
- sparse
- structured

Compressor Constraints



$$q_1 - q_2 + z_1 \cdot c_1(p_1, q_1, p_2, q_2) \geq 0$$

where p_i pressures

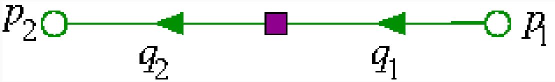
q_i flows

z_i 0–1 variables

$= 1$ if machine is on

c_i nonlinear functions

Compressor Constraints



$$q_1 - q_2 + z_1 \cdot c_1(p_1, q_1, p_2, q_2) \geq 0$$

where p_i pressures

q_i flows

z_i 0–1 variables

$= 1$ if machine is on

c_i nonlinear functions

In general:

$$A\,\mathbf{q} + \text{diag}(z_i)\,c(\mathbf{p}, \mathbf{q}) \geq \mathbf{0}$$

- non-linear
- sparse
- structured
- 0–1 variables

Other Constraints

Bounds on pressures and flows

$$p_{\min} \leq p \leq p_{\max}$$
$$q_{\min} \leq q \leq q_{\max}$$

- **In general:** Simple bounds on variables

Objectives

Many possible objectives

- maximize / minimize sum of pressures
- minimize compressor fuel cost
- minimize supply

Objectives

Many possible objectives

- maximize / minimize sum of pressures
- minimize compressor fuel cost
- minimize supply

+ combinations of these

Actual Data

British Gas National Transmission System

- 199 nodes
- 196 pipes
- 21 compressors

Actual Data

British Gas National Transmission System

- 199 nodes
- 196 pipes
- 21 compressors

Steady state problem

$\sim$ **400 variables**

Actual Data

British Gas National Transmission System

- 199 nodes
- 196 pipes
- 21 compressors

Steady state problem

$\sim$ **400 variables**

24-hour variable demand problem with 10 minute discretization

$\sim$ **58,000 variables**

Challenge: Solve this in real time!

Motivation for Course

This problem is **typical** of **real-world, large-scale** applications:

- linear constraints
- nonlinear constraints
- simple bounds
- structure
- integer variables
- global minimum “required”
- discretization

(Some) Other Application Areas

- minimum energy problems
- structural design problems
- traffic equilibrium models
- production scheduling problems
- portfolio selection
- parameter determination in financial markets
- hydro-electric power scheduling
- gas production models
- efficient models of alternative energy sources

Data fitting & inverse problems

An experiment is described by the nonlinear relation

$$y = F(p, x)$$

and repeated M times to get data pairs $(x_i, y_i)_{i=1}^{M}$.

Find parameters p that best fit our observations,

$$\min_{p \in \mathbb{R}^P} \sum_{i=1}^{M} |F(p, x_i) - y_i|^2$$

Input/Observation data $\quad (x_i, y_i)_{i=1}^{M}, \quad x_i \in \mathbb{R}^{K_I}, \quad y_i \in \mathbb{R}^{K_O}$

Forward map $F : \mathbb{R}^P \times \mathbb{R}^{K_I} :\to \mathbb{R}^{K_O}$

Parameters $p \in \mathbb{R}^P$

Data fitting & inverse problems

An experiment is described by the nonlinear relation

$$y = F(p, x)$$

and repeated M times to get data pairs $(x_i, y_i)_{i=1}^{M}$.

Find parameters p that best fit our observations,

$$\min_{p \in \mathbb{R}^P} \sum_{i=1}^{M} |F(p, x_i) - y_i|^2 + R(p)$$

Input/Observation data $(x_i, y_i)_{i=1}^{M}, \quad x_i \in \mathbb{R}^{K_I}, \quad y_i \in \mathbb{R}^{K_O}$

Forward map $F : \mathbb{R}^P \times \mathbb{R}^{K_I} :\to \mathbb{R}^{K_O}$

Parameters $p \in \mathbb{R}^P$

Regularization $R : \mathbb{R}^P \to \mathbb{R}_+$.

Examples of data fitting

- Back Propagation in Neural Networks

$$y = F(w, b, x),$$

where w are weights, b are biases and $(x_i, y_i)_{i=1}^M$ is training data. Training $\equiv$ find minimizing $p = (w, b)$.

- Computer Vision: 3D geometry and camera pose estimation based on photographs from multiple angles,

$$\min_{c,p} \sum_{i=1}^{C} \sum_{j=1}^{P} v_{i,j} |F(c_i, p_j) - y_{i,j}|^2.$$

Here p_j are 3D points. They are projected onto ith image at coordinates $F(c_i, p_j)$ provided $v_{i,j} = 1$, i.e. the point is visible. c_i are camera parameters.

Optimization over function spaces

- **Image processing**
  Given a noisy image, $f : \Omega \to \mathbb{R}$,
  compute the denoised image u,

$$\operatorname*{argmin}_{u \in \mathrm{BV}(\Omega)} \frac{\lambda}{2} \int_\Omega (f - u)^2 dx + \|u\|_{\mathrm{TV}(\Omega)} .$$

- **Optimal control as PDE constrained optimization**

$$u'(t) = A(p(t))\, u(t), \quad u(0) = u_0.$$

  Find the optimal control $p(t)$ such that at time T we are close to target state u_{target},

$$\operatorname*{argmin}_{\substack{p \in C^\infty([0,T]) \\ u' = A(p) u}} \left\| u(T) - u_{\text{target}} \right\|_2 .$$

Mathematical Problem Statement

Objective function $f : \mathbb{R}^N \to \mathbb{R}$ (smooth)

Admissible (or feasible) set

$$\Omega = \left\{\mathbf{x} \in \mathbb{R}^N : c_j(\mathbf{x}) = 0,\ j \in \mathcal{E},\ c_j(\mathbf{x}) \geq 0,\ j \in \mathcal{I}\right\},$$

with $\mathbf{c} : \mathbb{R}^N \to \mathbb{R}^{M_e+M_i}$, $\mathcal{E} = \{1, \ldots, M_e\}$, $\mathcal{I} = \{M_e+1, \ldots, M_e+M_i\}$

$\mathbf{x} \in \Omega$ is called *feasible* or *admissible*.

Mathematical Problem Statement

Objective function $f : \mathbb{R}^N \to \mathbb{R}$ (smooth)

Admissible (or feasible) set

$$\Omega = \left\{\mathbf{x} \in \mathbb{R}^N : c_j(\mathbf{x}) = 0,\ j \in \mathcal{E},\ c_j(\mathbf{x}) \geq 0,\ j \in \mathcal{I}\right\},$$

with $\mathbf{c} : \mathbb{R}^N \to \mathbb{R}^{M_e+M_i}$, $\mathcal{E} = \{1, \ldots, M_e\}$, $\mathcal{I} = \{M_e+1, \ldots, M_e+M_i\}$

$\mathbf{x} \in \Omega$ is called *feasible* or *admissible*.

$\mathbf{x}_* \in \Omega$ is a *global minimizer* of f in Ω if

$$f(\mathbf{x}_*) \leq f(\mathbf{x}) \qquad \forall \mathbf{x} \in \Omega. \tag{1}$$

Mathematical Problem Statement

Objective function $f : \mathbb{R}^N \to \mathbb{R}$ (smooth)

Admissible (or feasible) set

$$\Omega = \left\{\mathbf{x} \in \mathbb{R}^N : c_j(\mathbf{x}) = 0,\ j \in \mathcal{E},\ c_j(\mathbf{x}) \geq 0,\ j \in \mathcal{I}\right\},$$

with $\mathbf{c} : \mathbb{R}^N \to \mathbb{R}^{M_e+M_i}$, $\mathcal{E} = \{1, \dots, M_e\}$, $\mathcal{I} = \{M_e+1, \dots, M_e+M_i\}$

$\mathbf{x} \in \Omega$ is called *feasible* or *admissible*.

$\mathbf{x}_* \in \Omega$ is a *global minimizer* of f in Ω if

$$f(\mathbf{x}_*) \leq f(\mathbf{x}) \qquad \forall \mathbf{x} \in \Omega. \tag{1}$$

$\mathbf{x}_* \in \Omega$ is a *(strict) local minimizer* of f in Ω if $\exists r > 0$ s.t.

$$f(\mathbf{x}_*) \leq f(\mathbf{x}) \qquad \forall \mathbf{x} \in \Omega \cap B_r(\mathbf{x}_*) \tag{2}$$

$$f(\mathbf{x}_*) < f(\mathbf{x}) \qquad \forall \mathbf{x} \in (\Omega \cap B_r(\mathbf{x}_*)) \setminus \{\mathbf{x}_*\} \quad \text{(strict)} \tag{3}$$

Mathematical Problem Statement

Objective function $f : \mathbb{R}^N \to \mathbb{R}$ (smooth)

Admissible (or feasible) set

$$\Omega = \left\{ \mathbf{x} \in \mathbb{R}^N : c_j(\mathbf{x}) = 0,\ j \in \mathcal{E},\ c_j(\mathbf{x}) \geq 0,\ j \in \mathcal{I} \right\},$$

with $\mathbf{c} : \mathbb{R}^N \to \mathbb{R}^{M_e+M_i}$, $\mathcal{E} = \{1, \dots, M_e\}$, $\mathcal{I} = \{M_e+1, \dots, M_e+M_i\}$

$\mathbf{x} \in \Omega$ is called *feasible* or *admissible*.

$\mathbf{x}_* \in \Omega$ is a *global minimizer* of f in Ω if

Hard unless convex!

$$f(\mathbf{x}_*) \leq f(\mathbf{x}) \qquad \forall \mathbf{x} \in \Omega. \tag{1}$$

$\mathbf{x}_* \in \Omega$ is a *(strict) local minimizer* of f in Ω if $\exists r > 0$ s.t.

Our focus

$$f(\mathbf{x}_*) \leq f(\mathbf{x}) \qquad \forall \mathbf{x} \in \Omega \cap B_r(\mathbf{x}_*) \tag{2}$$

$$f(\mathbf{x}_*) < f(\mathbf{x}) \qquad \forall \mathbf{x} \in (\Omega \cap B_r(\mathbf{x}_*)) \setminus \{\mathbf{x}_*\} \quad \text{(strict)} \tag{3}$$

Problems to be Considered

Constrained optimisation

$$\min_{\mathbf{x}\in\Omega\subset\mathbb{R}^N} f : \mathbb{R}^N \to \mathbb{R}$$

Problems to be Considered

Constrained optimisation

$$\min_{\mathbf{x} \in \Omega \subset \mathbb{R}^N} f : \mathbb{R}^N \to \mathbb{R}$$

Unconstrained optimisation

$$\min_{\mathbf{x} \in \mathbb{R}^N} f : \mathbb{R}^N \to \mathbb{R}$$

Problems to be Considered

Constrained optimisation

$$\min_{\mathbf{x}\in\Omega\subset\mathbb{R}^N} f : \mathbb{R}^N \to \mathbb{R}$$

Unconstrained optimisation

$$\min_{\mathbf{x}\in\mathbb{R}^N} f : \mathbb{R}^N \to \mathbb{R}$$

Solving simultaneous nonlinear equations

$$\text{Find} \quad \mathbf{x}_* \in \mathbb{R}^N \quad s.t. \quad F(\mathbf{x}_*) = \mathbf{0} \quad \text{for} \quad F : \mathbb{R}^N \to \mathbb{R}^N$$

Problems to be Considered

Constrained optimisation

$$\min_{\mathbf{x}\in\Omega\subset\mathbb{R}^N} f : \mathbb{R}^N \to \mathbb{R}$$

Unconstrained optimisation

$$\min_{\mathbf{x}\in\mathbb{R}^N} f : \mathbb{R}^N \to \mathbb{R}$$

Solving simultaneous nonlinear equations

$$\text{Find } \ \mathbf{x}_* \in \mathbb{R}^N \ \ s.t. \quad F(\mathbf{x}_*) = \mathbf{0} \quad \text{for} \quad F : \mathbb{R}^N \to \mathbb{R}^N$$

In reverse order!!

Course Overview

- Revision: Linear Algebra & Multivariate Calculus
- Newton's Method
- Global Convergence: Line Search & Trust Region
- Quasi-Newton Methods (gradient-free)
- Optimality Conditions for Constrained Optimisation
- Penalty Methods
- Barrier Methods
- Large-Scale Optimisation

Literature

- **Christoph Ortner**, *Continuous Optimization*, Lecture Notes, Oxford, 2009
- JE Dennis & RB Schnabel, *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*, 1983
- J Nocedal & SJ Wright, *Numerical Optimization*, 2006
- N.I.M. Gould & S. Leyffer, "An introduction to algorithms for nonlinear optimization", in *Frontiers in NA*, Springer, 2003

Other Important Informations (see also handout)

Lectures:	Wednesday	10.15	1W 3.30
	Thursday	11.15	CB 4.10
No lectures in	week 4		
Problem classes:	Friday	17.15	8W 2.20
No problem classes in	weeks 1,2,4,10		

- Problem sheets, handouts, Matlab codes, and other useful material/links available on the course web page

  `http://www.pranavsingh.co.uk/ma40050`.

- Computing in **Matlab** (on BUCS or own laptop/PC).

- **Assignment** (worth 25%) provisionally planned Mar 18–Apr 28.