Nonlinear programming

Computational Methods in Optimization

CS 520, Purdue

David F. Gleich

Purdue University

April 20, 2017

Problems

Equality constrained

$$\begin{array}{ll} \text{minimize} & f(\mathbf{x}) \\ \text{subject to} & \mathbf{c}(\mathbf{x}) = 0 \end{array}$$

Inequality constrained

$$\begin{array}{ll} \text{minimize} & f(\mathbf{x}) \\ \text{subject to} & \mathbf{d}(\mathbf{x}) \geq 0 \end{array}$$

General optimization

$$\begin{array}{ll} \text{minimize} & f(\mathbf{x}) \\ \text{subject to} & \mathbf{c}(\mathbf{x}) = 0 \\ & \mathbf{d}(\mathbf{x}) \geq 0 \end{array}$$

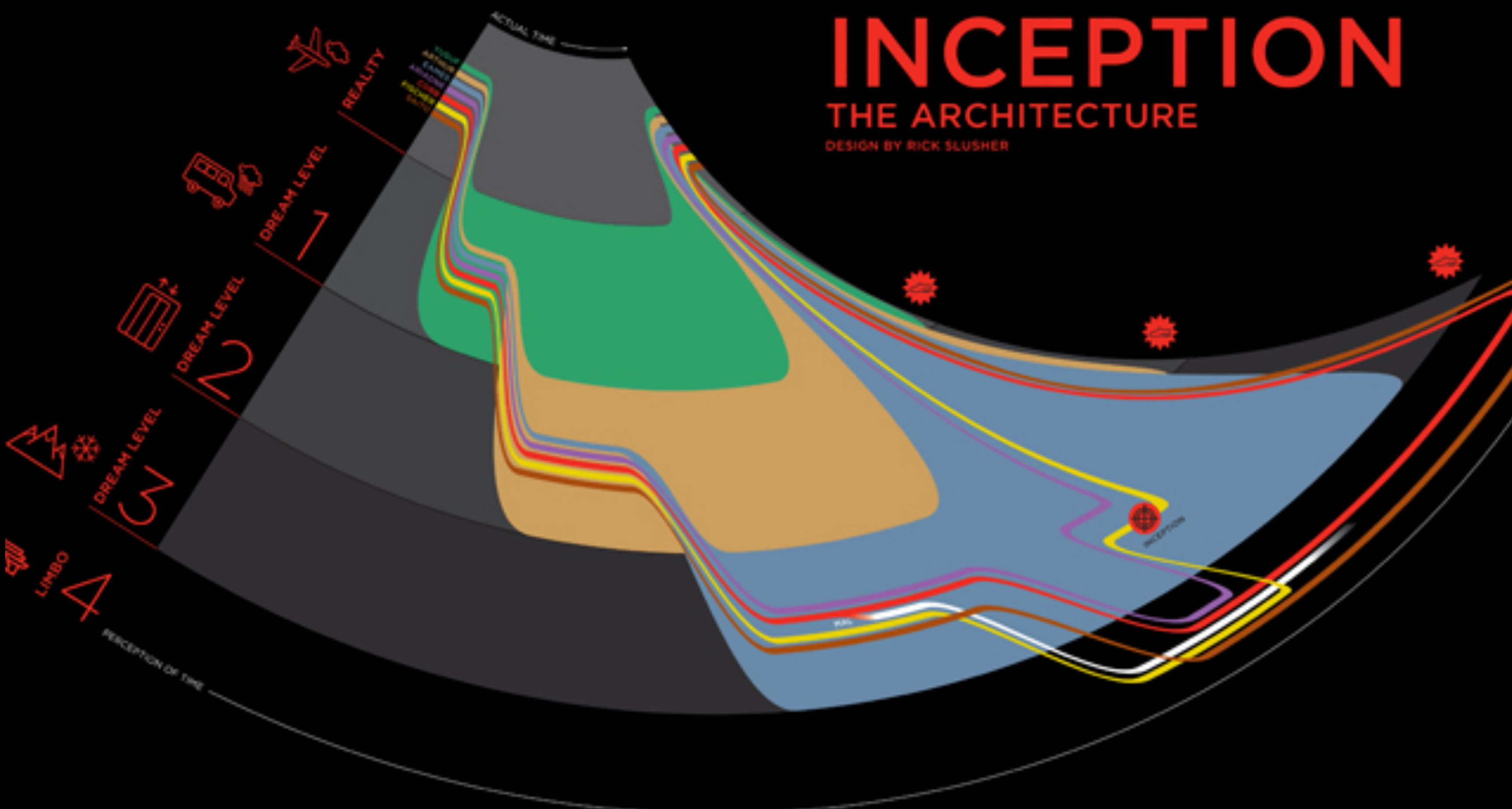
Overarching idea

Approximate these problems by something

- easier, or
- more simple

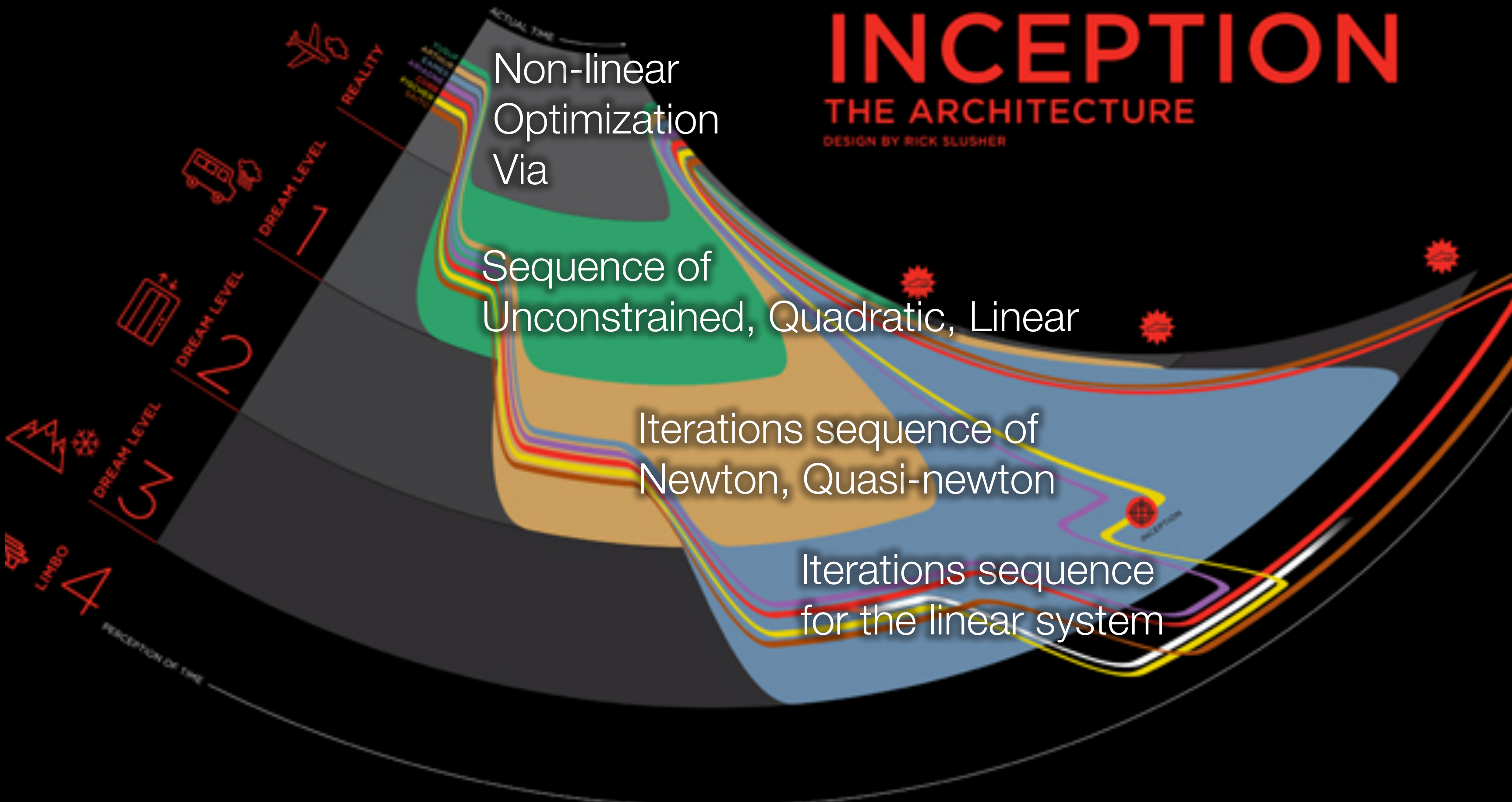
And then solve a sequence of optimization problems.

Inception



INCEPTION
THE ARCHITECTURE
DESIGN BY RICK SLUSHER
ACTUAL TIME
REALITY
DREAM LEVEL
1
DREAM LEVEL
2
DREAM LEVEL
3
LIMBO
4
INCEPTION
PERCEPTION OF TIME

Inception


INCEPTION
THE ARCHITECTURE
DESIGN BY RICK SLUSHER
REALITY
DREAM LEVEL 1
DREAM LEVEL 2
DREAM LEVEL 3
LIMBO 4
Non-linear
Optimization
Via
Sequence of
Unconstrained, Quadratic, Linear
Iterations sequence of
Newton, Quasi-newton
Iterations sequence
for the linear system

The equality constrained problem

$$\begin{array}{ll} \text{minimize} & f(\mathbf{x}) \\ \text{subject to} & \mathbf{c}(\mathbf{x}) = 0 \end{array}$$

Chapter 17, Nocedal & Wright

Penalty Methods and
Augmented Lagrangians

A penalty method

$$\begin{aligned} &\text{minimize} \quad f(\mathbf{x}) \\ &\text{subject to} \quad \mathbf{c}(\mathbf{x}) = 0 \end{aligned}$$

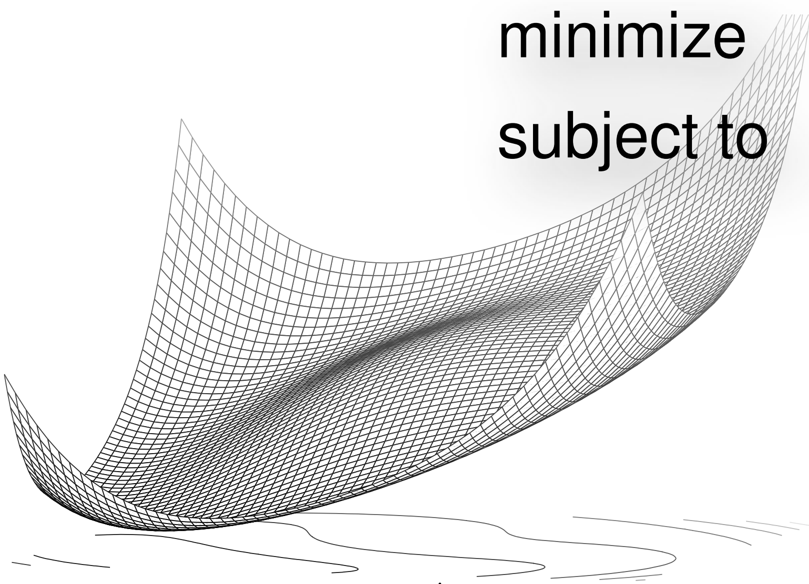
$$\text{minimize} \quad f(\mathbf{x}) + \mu \textstyle\sum_i c_i(\mathbf{x})^2 \quad \leftrightarrow \quad \text{minimize} \quad f(\mathbf{x}) + \mu \mathbf{c}(\mathbf{x})^T \mathbf{c}(\mathbf{x})$$

A penalty method

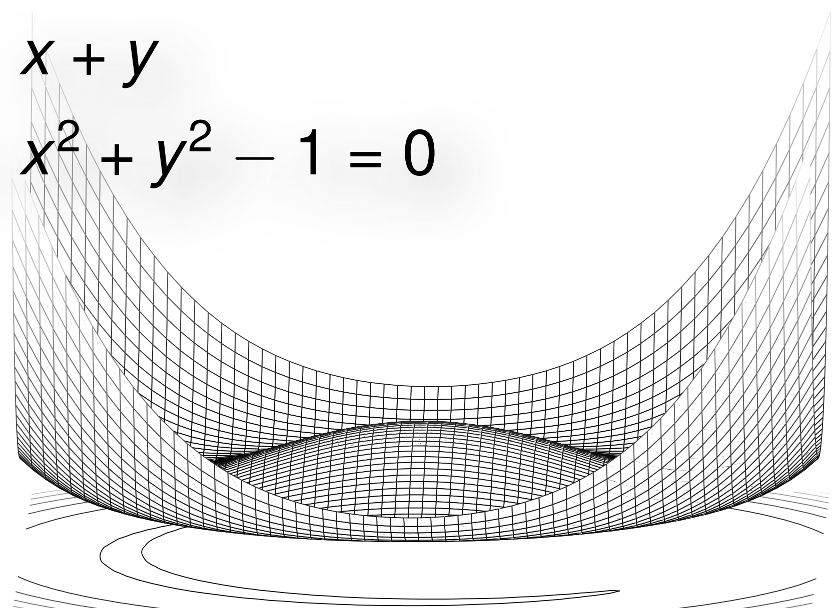
$$\begin{array}{ll} \text{minimize} & f(\mathbf{x}) \\ \text{subject to} & \mathbf{c}(\mathbf{x}) = 0 \end{array}$$

$$\text{minimize} \quad f(\mathbf{x}) + \mu \textstyle\sum_i c_i(\mathbf{x})^2 \quad \leftrightarrow \quad \text{minimize} \quad f(\mathbf{x}) + \mu \mathbf{c}(\mathbf{x})^T \mathbf{c}(\mathbf{x})$$

$$\begin{array}{ll} \text{minimize} & x + y \\ \text{subject to} & x^2 + y^2 - 1 = 0 \end{array}$$



$\mu = 1/2$



$\mu = 50$

A penalty method

$$\begin{array}{ll} \text{minimize} & f(\mathbf{x}) \\ \text{subject to} & \mathbf{c}(\mathbf{x}) = 0 \end{array}$$

$$\text{minimize} \quad f(\mathbf{x}) + \mu \textstyle\sum_i c_i(\mathbf{x})^2 \;\; \leftrightarrow \;\; \text{minimize} \quad f(\mathbf{x}) + \mu \mathbf{c}(\mathbf{x})^T \mathbf{c}(\mathbf{x})$$

Let $\{\tau_k\} \to 0$, $\{\mu_k\} \to \infty$.

While $\|\mathbf{c}(\mathbf{x}_k)\| \geq \text{tol}$

Solve $\text{minimize} \quad f(\mathbf{x}) + \mu_k/2\mathbf{c}(\mathbf{x})^T\mathbf{c}(\mathbf{x})$ to tolerance τ_k (Gradient norm)

Set $\mathbf{x}^{(k+1)}$ to be the solution.

Convergence of the penalty method

If we'll be able to prove this convergences, we'll need a strong condition.

Why?

Convergence of the penalty method

If we'll be able to prove this convergences, we'll need a strong condition.

$$\mathbf{c}(\mathbf{x})^T\mathbf{c}(\mathbf{x}) \text{ vs. } \mathbf{c}(\mathbf{x}) = 0$$

Convergence of the penalty method

If we'll be able to prove this convergences, we'll need a strong condition.

$$\mathbf{c}(\mathbf{x})^T\mathbf{c}(\mathbf{x}) \text{ vs. } \mathbf{c}(\mathbf{x}) = 0$$

Theorem 17.1 (Paraphrased)

If we use the global minimizer of each subproblem, then we solve the problem in the $\mu_k \to \infty$ limit

Convergence of the penalty method

Theorem 17.2 (Paraphrased)

If we approximately minimize each problem to a point where

$$\|\mathbf{g}(\mathbf{x}_k)\| \leq \tau_k$$

Then either a limit point of the sequence is either (infeasible) and a stationary point of $\|\mathbf{c}(\mathbf{x})\|^2$

Or

It's a KKT point of the original problem

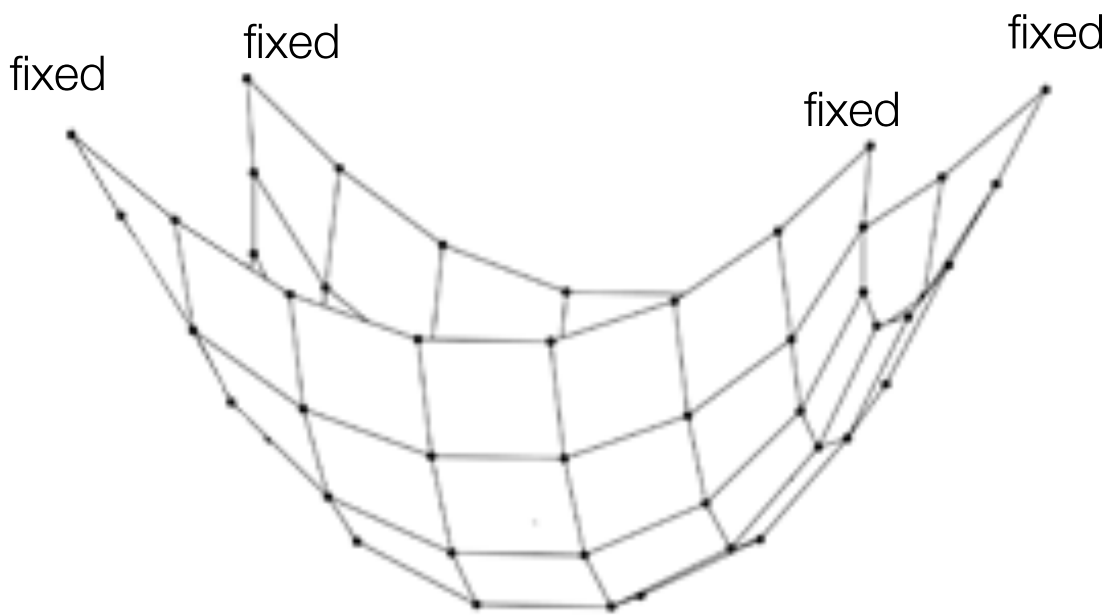
Weaknesses

Highly "ill-conditioned" as $\mu_k \to \infty$

Convergence only about limit points

Problems with penalty methods

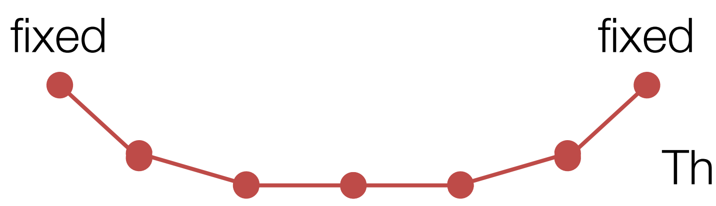
The hanging net problem



minimize the energy in the system
subject to using “steel” links

Problems with penalty methods

The hanging net problem



These links would stretch! (Small constraint violation gives a large change in the energy objective)
As $\mu_k \to \infty$ we'd violate these constraints "more" than others.

minimize the energy in the system
subject to using "steel" links

Problems with penalty methods

All constraints are not equal!

Augmented Lagrangian Methods

$$\begin{array}{ll} \text{minimize} & f(\mathbf{x}) \\ \text{subject to} & \mathbf{c}(\mathbf{x}) = 0 \end{array}$$

$$\mathcal{L}(\mathbf{x}; \lambda, \mu) = f(\mathbf{x}) - \lambda^T \mathbf{c}(\mathbf{x}) + \mu/2 \|\mathbf{c}(\mathbf{x})\|^2.$$

Augmented Lagrangian Methods

$$\begin{aligned} \text{minimize} \quad & f(\mathbf{x}) \\ \text{subject to} \quad & \mathbf{c}(\mathbf{x}) = 0 \end{aligned}$$

$$\mathcal{L}(\mathbf{x}; \lambda, \mu) = f(\mathbf{x}) - \lambda^T \mathbf{c}(\mathbf{x}) + \mu/2 \|\mathbf{c}(\mathbf{x})\|^2.$$

If we minimize in x alone

$$\nabla \mathcal{L}(\mathbf{x}) = \mathbf{g}_f(\mathbf{x}) - \mathbf{J}_c(\mathbf{x})^T (\lambda - \mu \mathbf{c}(\mathbf{x})) = 0$$

One of the KKT conditions of the non-linear program is

$$\mathbf{g}_f(\mathbf{x}^*) - \mathbf{J}_c(\mathbf{x}^*)^T \lambda^* = 0$$

Augmented Lagrangian Methods

$$\nabla \mathcal{L}(\mathbf{x}) = \mathbf{g}_f(\mathbf{x}) - \mathbf{J}_c(\mathbf{x})^T(\lambda - \mu \mathbf{c}(\mathbf{x})) = 0$$

$$\mathbf{g}_f(\mathbf{x}^*) - \mathbf{J}_c(\mathbf{x}^*)^T \lambda^* = 0$$

So in an algorithm, we use

$$\lambda_{k+1} = \lambda_k - \mu_k \mathbf{c}(\mathbf{x})$$

$\underset{\mathbf{x}}{\text{minimize}} \quad \mathcal{L}(\mathbf{x}; \lambda_k, \mu_k)$ to tolerance τ_k starting from $\mathbf{x}_k$
if $\|\mathbf{c}(\mathbf{x})\|$ is small, stop!
else set $\lambda_{k+1} = \lambda_k - \mu_k \mathbf{c}(\mathbf{x}_k)$
set $\mu_{k+1} \geq \mu_k$

Convergence of AL methods

See theorem 17.5 and 17.6

See Alg 17.4 for the method used in LANCELOT with bound-constraints.

Barrier methods

$$\begin{array}{ll} \text{minimize} & f(\mathbf{x}) \\ \text{subject to} & \mathbf{d}(\mathbf{x}) \geq 0 \end{array}$$

$$\text{minimize} \quad f(\mathbf{x}) - \mu \mathbf{e}^T \log(\mathbf{d}(\mathbf{x}))$$

Chapter 16, Griva, Sofer & Nash

The inequality constrained problem

$$\begin{aligned} \text{minimize} \quad & f(\mathbf{x}) \\ \text{subject to} \quad & \mathbf{c}(\mathbf{x}) = 0 \\ & \mathbf{d}(\mathbf{x}) \geq 0 \end{aligned}$$

can be transformed into

$$\begin{aligned} \text{minimize} \quad & f(\mathbf{x}) \\ \text{subject to} \quad & \mathbf{c}(\mathbf{x}) = 0 \\ & \mathbf{d}(\mathbf{x}) - \mathbf{s} = 0 \\ & \mathbf{l} \leq \mathbf{x} \leq \mathbf{u}, \mathbf{s} \geq 0 \end{aligned}$$

So handling equality, and bounds suffices!

The bound constrained problem

$$\begin{aligned} \text{minimize} \quad & f(\mathbf{x}) \\ \text{subject to} \quad & \mathbf{c}(\mathbf{x}) = 0 \\ & \mathbf{l} \leq \mathbf{x} \leq \mathbf{u} \end{aligned}$$

See algorithms in
Algorithm 17.4, Chapter 18
Nocedal & Wright

LANCELOT
Sequential quadratic programming
Gradient projection