Module 1: Formulations (Nonlinear Models)

### So far ...

• Linear programs, and

 $\min c^T x$ <br/>s.t.  $Ax \ge b$ <br/> $x \ge 0$ 

### So far ...

- Linear programs, and
- Integer linear programs.

Both have linear/affine constraints.

```
\begin{array}{l} \min \ c^T x \\ \text{s.t.} \ Ax \geq b \\ x \geq 0 \\ x \ \text{integer} \end{array}
```

## So far ...

- Linear programs, and
- Integer linear programs.

Both have linear/affine constraints.

Now: Nonlinear generalization!

 $\begin{array}{l} \min \ c^T x \\ \text{s.t.} \ Ax \geq b \\ x \geq 0 \\ x \ \text{integer} \end{array}$ 

A non-linear program (NLP) is of the form

min 
$$f(x)$$
  
s.t.  $g_1(x) \le 0$   
 $g_2(x) \le 0$   
...  
 $g_m(x) \le 0,$ 

A non-linear program (NLP) is of the form

where

• 
$$x \in \mathbb{R}^n$$
,

min 
$$f(x)$$
  
s.t.  $g_1(x) \le 0$   
 $g_2(x) \le 0$   
...  
 $g_m(x) \le 0$ ,

A non-linear program (NLP) is of the form

 $- \mathbb{D}^n$ 

where

• 
$$x \in \mathbb{R}^n$$
,

• 
$$f: \mathbb{R}^n \to \mathbb{R}$$
, and

$$\begin{array}{ll} \min & f(x) \\ \text{s.t.} & g_1(x) \leq 0 \\ & g_2(x) \leq 0 \\ & \dots \\ & g_m(x) \leq 0, \end{array}$$

A non-linear program (NLP) is of the form

 $\begin{array}{ll} \min & f(x) \\ \text{s.t.} & g_1(x) \leq 0 \\ & g_2(x) \leq 0 \\ & \dots \end{array}$ 

 $g_m(x) \le 0,$ 

where

- $x \in \mathbb{R}^n$ ,
- $f: \mathbb{R}^n \to \mathbb{R}$ , and

• 
$$g_i : \mathbb{R}^n \to \mathbb{R}$$
.

A non-linear program (NLP) is of the form

min f(x)s.t.  $g_1(x) \le 0$  $g_2(x) \le 0$ ...  $g_m(x) \le 0,$  where

• 
$$x \in \mathbb{R}^n$$
,

•  $f:\mathbb{R}^n \rightarrow \, \mathbb{R}$  , and

• 
$$g_i : \mathbb{R}^n \to \mathbb{R}$$
.

Note: Linear programs are NLPs!

#### Example 1: Finding Close Points in an LP

# Finding Close Points in an LP

Problem: we are given an LP (P), and an infeasible  $\min c^T x$ point  $\bar{x}$ . s.t.  $x \in P$ 

Goal: find a point  $x \in P$  that is as close as possible to  $\bar{x}$ .

 $\begin{array}{ll} \min \ c^T x \\ \text{s.t.} \ x \in P \end{array}$ 

Goal: find a point  $x \in P$  that is as close as possible to  $\bar{x}$ .

e.g.: find a point  $x \in P$  that minimizes the Euclidean distance to  $\bar{x}$ :

$$||x - \bar{x}||_2 = \sqrt{\sum_{i=1}^n (x_i - \bar{x}_i)^2}$$

 $\begin{array}{ll} \min \ c^T x \\ \text{s.t.} \ x \in P \end{array}$ 

Goal: find a point  $x \in P$  that is as close as possible to  $\bar{x}$ .

e.g.: find a point  $x \in P$  that minimizes the Euclidean distance to  $\bar{x}$ :

$$||x - \bar{x}||_2 = \sqrt{\sum_{i=1}^{n} (x_i - \bar{x}_i)^2}$$

**Remark**:  $||p||_2$  is called the  $L^2$ -norm of p

 $\begin{array}{ll} \min \ c^T x \\ \text{s.t.} \ x \in P \end{array}$ 

Goal: find a point  $x \in P$  that is as close as possible to  $\bar{x}$ .

e.g.: find a point  $x \in P$  that minimizes the Euclidean distance to  $\bar{x}$ :

$$||x - \bar{x}||_2 = \sqrt{\sum_{i=1}^{n} (x_i - \bar{x}_i)^2}$$

**Remark**:  $||p||_2$  is called the  $L^2$ -norm of p

 $\begin{array}{ll} \min \ c^T x \\ \text{s.t.} \ x \in P \end{array}$ 

$$P = \{x : Ax \le b\}$$

$$\min \|x - \bar{x}\|_2$$
  
s.t.  $x \in P$ 

#### Example 2: Binary IP via NLP

Suppose we are given a binary IP (i.e., an integer program all of whose variables are binary).

 $\begin{array}{ll} \max \ c^T x\\ \text{s.t.} \ Ax \leq b\\ x \geq 0\\ x_j \in \{0,1\} \quad (j \in \{1,\ldots,n\}) \end{array}$ 

Suppose we are given a binary IP (i.e., an integer program all of whose variables are binary).

Recall: (binary) IPs are generally hard to solve!

 $\begin{array}{ll} \max \ c^T x \\ \text{s.t.} \ Ax \leq b \\ x \geq 0 \\ x_j \in \{0,1\} \quad (j \in \{1,\ldots,n\}) \end{array}$ 

Suppose we are given a binary IP (i.e., an integer program all of whose variables are binary).

 $\begin{array}{ll} \max \ c^T x\\ \text{s.t.} \ Ax \leq b\\ x \geq 0\\ x_j \in \{0,1\} \quad (j \in \{1,\ldots,n\}) \end{array}$ 

Recall: (binary) IPs are generally hard to solve!

Now: can write any binary IP as an NLP!

Suppose we are given a binary IP (i.e., an integer program all of whose variables are binary).

Recall: (binary) IPs are generally hard to solve!

Now: can write any binary IP as an NLP!

Ideas?

$$\begin{array}{ll} \max \ c^T x \\ \text{s.t.} \ Ax \leq b \\ x \geq 0 \\ x_i \in \{0,1\} \quad (j \in \{1,\ldots,n\}) \end{array}$$

Suppose we are given a binary IP (i.e., an integer program all of whose variables are binary).

Now: can write any binary IP as an NLP!

Ideas?

$$\begin{array}{ll} \max \ c^T x \\ \text{s.t.} \ Ax \leq b \\ x \geq \mathbb{O} \\ x_j(1-x_j) = 0 \quad (j \in [n]) \quad (\star) \end{array}$$

Suppose we are given a binary IP (i.e., an integer program all of whose variables are binary).

Recall: (binary) IPs are generally hard to solve!

Now: can write any binary IP as an NLP!

$$\begin{array}{ll} \max \ c^T x \\ \text{s.t.} \ Ax \leq b \\ x \geq 0 \\ x_j(1-x_j) = 0 \quad (j \in [n]) \quad (\star) \end{array}$$

Correctness: For  $j \in [n]$ , (\*) is holds iff  $x_j = 0$  or  $x_j = 1$ .

Ideas?

Suppose we are given a binary IP (i.e., an integer program all of whose variables are binary).

Recall: (binary) IPs are generally hard to solve!

Now: can write any binary IP as an NLP!

Ideas?

$$\begin{array}{ll} \max \ c^T x \\ \text{s.t.} \ Ax \leq b \\ x \geq \mathbb{O} \\ x_j(1-x_j) = 0 \quad (j \in [n]) \quad (\star) \end{array}$$

Correctness: For  $j \in [n]$ , (\*) is holds iff  $x_j = 0$  or  $x_j = 1$ .

Q: Can you change the NLP to express the fact that  $x_j$  is any non-negative integer instead of binary?

Suppose we are given a binary IP (i.e., an integer program all of whose variables are binary).

Recall: (binary) IPs are generally hard to solve!

Now: can write any binary IP as an NLP!

Ideas?

$$\begin{array}{ll} \max \ c^T x \\ \text{s.t.} \ Ax \leq b \\ x \geq 0 \\ \sin(\pi \, x_j) = 0 \quad (j \in [n]) \quad (*) \end{array}$$

Correctness: For  $j \in [n]$ , (\*) is holds iff  $x_j = 0$  or  $x_j = 1$ .

Q: Can you change the NLP to express the fact that  $x_j$  is any non-negative integer instead of binary?

Correctness: note that  $sin(\pi x_j) = 0$  only if  $x_j$  is an integer.

#### Example 3: Fermat's Last Theorem

#### **Conjecture** [Fermat, 1637] There are no integers $x, y, z \ge 1$ and $n \ge 3$ such

that

$$x^n + y^n = z^n.$$



#### Conjecture [Fermat, 1637]

There are no integers  $x, y, z \ge 1$  and  $n \ge 3$  such that

$$x^n + y^n = z^n.$$



Arithmeticoum Liber II. Minimum, animum and and the sign offer in the sign offer in the sign of the sign offer in the sign offer in the sign of the sign of the sign offer in the sign of the

#### W STATTIONEN FIL

EXTINGENT Spectra value of par A spectra presided gardiani, al can al aspect coins to pedaran menufic measurable manufacturation pattern in: coins become interference or manifulant eff.

QVASTIO VIIL

Description of the second seco	the nearby in any second to it and

#### OBSERVATIO DOMINI PETRI DE FERMAT.

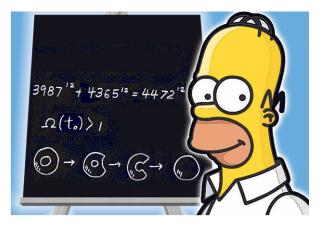
Crime anter in dee sales, our productspedentes in deer productspedentes de gestreleer militer in opfisiene idea geschenen perform in dee sing-ten menter for of desider meter ou demosfracteene medicien free densi-den menter en opfisier en anter ou

#### QUASTIO IN.

where $d$ - mapping the set of t

This is false ...

#### This is false ...



... doh!

#### **Conjecture** [Fermat, 1637] There are no integers $x, y, z \ge 1$ and $n \ge 3$ such

that

$$x^n + y^n = z^n.$$



#### Conjecture [Fermat, 1637]

There are no integers  $x,\,y,\,z\geq 1$  and  $n\geq 3$  such that

$$x^n + y^n = z^n.$$

In the margins of a copy of a 1670 article of Diophantus Arithmetica he wrote that he had a proof that was a bit too large to fit.



Anthmetico	um Liber II. 6
introducts transverses, minor annex (N. sept Movement (N. sept Movement (N. sept Movement (N. sept Movement (N. september eff)) and the first definition of the foregoing of the set of the se	$ \begin{array}{l} f(h) = i  h_{0}  \operatorname{col} \left[ f(h)  h(h) + f(h)  h(h) \right] \\ & = h_{0}  \operatorname{col} \left[ h(h) + f(h)  h(h) \right] \\ & = h_{0}  \operatorname{col} \left[ h(h) + f(h)  h(h) \right] \\ & = h_{0}  \operatorname{col} \left[ h(h) + f(h)  h(h) \right] \\ & = h_{0}  \operatorname{col} \left[ h(h) + f(h)  h(h) \right] \\ & = h_{0}  \operatorname{col} \left[ h(h) + f(h)  h(h) \right] \\ & = h_{0}  \operatorname{col} \left[ h(h) + f(h)  h(h) \right] \\ & = h_{0}  \operatorname{col} \left[ h(h) + f(h)  h(h) \right] \\ & = h_{0}  \operatorname{col} \left[ h(h)  h(h)  h(h) \right] \\ & = h_{0}  \operatorname{col} \left[ h(h)  h(h)  h(h) \right] \\ & = h_{0}  \operatorname{col} \left[ h(h)  h(h)  h(h) \right] \\ & = h_{0}  \operatorname{col} \left[ h(h)  h(h)  h(h)  h(h) \right] \\ & = h_{0}  \operatorname{col} \left[ h(h)  h(h)  h(h)  h(h) \right] \\ & = h_{0}  \operatorname{col} \left[ h(h)  h(h)  h(h)  h(h) \right] \\ & = h_{0}  \operatorname{col} \left[ h(h)  h(h)  h(h)  h(h)  h(h) \right] \\ & = h_{0}  \operatorname{col} \left[ h(h)  h(h)  h(h)  h(h)  h(h)  h(h) \right] \\ & = h_{0}  \operatorname{col} \left[ h(h)  h(h)  h(h)  h(h)  h(h)  h(h)  h(h) \right] \\ & = h_{0}  \operatorname{col} \left[ h(h)  h($

# STATTIONANY

Contra traphic plan is applied only main of par & oppole parameter quellant, all onin Contra traphic plan is quellarse menufications provide minoritements quellencin. A Contra talem in contra traine indepent, et maniplant eff.

<sup>1</sup> in these spectration, is repetitioned in weighting in the symptomeses, Forugation, and the symptomeses of the sympositic start if a second start of the symposite start if a second start of the symposite start if a second start is second start of the symposite start if a second start is second start in the symposite start if a second start is a second start in the symposite start if a second start is a second start of the symposite start if a second start is a second start in the symposite start if a second start is a second start of the symposite start is a	111 Wheelphills register a total of the second probability of the sec
--	---

#### ESERVATIO DOMINI PETRI DE FERMAT.

Crister attore in date sales, our production in date, production production date structure and the second structure performs in date shipdate structure for of divident structure dates/backman minimized for dates/ date structure for of divident structure dates/backman minimized for dates/ dates structure for of divident structure dates/ dates structure dates at the structure dates at t

#### 7.45T10 18.

where $d$ - mapping the set of t

# Conjecture [Fermat, 1637]

There are no integers  $x,\,y,\,z\geq 1$  and  $n\geq 3$  such that

$$x^n + y^n = z^n.$$

In the margins of a copy of a 1670 article of Diophantus Arithmetica he wrote that he had a proof that was a bit too large to fit.

Some 358 years later, Sir Andrew Wiles gave the first accepted proof of the theorem.



Anthmetico		61
introduces transverses, minute annual (N), mpar bloc minut (N, $\rightarrow$ ). Optimiz transverses (N, $\rightarrow$ ), dependent of the definition of the def	$ \begin{array}{l} f(h) \in I \mbox{ transmitted} \ for \ f \ h) \\ re \ for \ f$	F. la

W LIMITIONEN

а,	x	x	5	11	0	٧	11	L.	

$\begin{array}{llllllllllllllllllllllllllllllllllll$

#### ESERVATIO DOMINI PETRI DE FERMAT.

Criste anter ice des subse, our produces produces in dem produces the general-ter matter incorporation who quadwarm peopleses in desc objeter menter for of devices more net another homofenations models for description mergine exignate net append.

#### 45T10 18.

# Conjecture [Fermat, 1637]

There are no integers  $x,\,y,\,z\geq 1$  and  $n\geq 3$  such that

$$x^n + y^n = z^n.$$

In the margins of a copy of a 1670 article of Diophantus Arithmetica he wrote that he had a proof that was a bit too large to fit.

Some 358 years later, Sir Andrew Wiles gave the first accepted proof of the theorem. The proof is over 150 pages long!



Anthmetico	rum Liber II. 61
introduces consistent with the second secon	$ \begin{array}{l} c^{2} \log x + dy_{0} \mbox{ solution } \mathcal{F} \mbox{ solution } $
14 21M13	TANAN PIL

QVASTIO VIIL

Index quadratics, inspectrum for we hadron quadratics, in a quadratics, Promity guarances, Queges, and Queges, Quester, Queges, Quege	The three phase of the second
---	---

SERVATIO DOMINI PETRI DE FERMAT

Che parto actor teles alto, cue padonapadente inder protorspedente Che parto contexe militari recipitane teta padonam porfesse in dan objdan mentio for ef divider sono ner domefutatione metodolm fine dense, fan mentio for ef divider ante opera.

#### ASTI0 18.

R <sup>1</sup> /1111 operating approximation and distance of some productions, Post- mer networks prime trains (N. Arberter spat- serence), processing one devices of the analysis (N. Strains, Post- metric), and the source of the source of the prime strains (N. Strains, Post- metric), and the source of the source of the prime strains (N. Strains, P. Strains, Post- singurate strains on the Strain, N. Strain Strains, Strains, Strains, S. Strain, S. Strains, Strains, S. Strains, Strai	$ \begin{array}{c} \mathbf{E}^{(1)} \left( \boldsymbol{\beta} \right) & \text{where the } \boldsymbol{\beta} & wavely have a consideration of the set of th$
---	---

$$\begin{array}{ll} \min & (x_1^{x_4} + x_2^{x_4} - x_3^{x_4})^2 \\ & + (\sin \pi \, x_1)^2 + (\sin \pi \, x_2)^2 + (\sin \pi \, x_3)^2 + (\sin \pi \, x_4)^2 \\ \text{s.t.} & x_i \ge 1 \quad (i = 1 \dots 3) \\ & x_4 \ge 3 \end{array}$$

$$\begin{array}{ll} \min & (x_1^{x_4} + x_2^{x_4} - x_3^{x_4})^2 \\ & + (\sin \pi \, x_1)^2 + (\sin \pi \, x_2)^2 + (\sin \pi \, x_3)^2 + (\sin \pi \, x_4)^2 \\ \text{s.t.} & x_i \ge 1 \quad (i = 1 \dots 3) \\ & x_4 \ge 3 \end{array}$$

• The NLP is trivially feasible, and

$$\begin{array}{ll} \min & (x_1^{x_4} + x_2^{x_4} - x_3^{x_4})^2 \\ & + (\sin \pi \, x_1)^2 + (\sin \pi \, x_2)^2 + (\sin \pi \, x_3)^2 + (\sin \pi \, x_4)^2 \\ \text{s.t.} & x_i \ge 1 \quad (i = 1 \dots 3) \\ & x_4 \ge 3 \end{array}$$

- The NLP is trivially feasible, and
- the value of any feasible solution is non-negative as its objective is the sum of squares.

$$\begin{array}{ll} \min & (x_1^{x_4} + x_2^{x_4} - x_3^{x_4})^2 \\ & + (\sin \pi \, x_1)^2 + (\sin \pi \, x_2)^2 + (\sin \pi \, x_3)^2 + (\sin \pi \, x_4)^2 \\ \text{s.t.} & x_i \ge 1 \quad (i = 1 \dots 3) \\ & x_4 \ge 3 \end{array}$$

- The NLP is trivially feasible, and
- the value of any feasible solution is non-negative as its objective is the sum of squares.
- In fact, the value of a solution  $(x_1, x_2, x_3, x_4)$  is 0 iff

$$\min (x_1^{x_4} + x_2^{x_4} - x_3^{x_4})^2 + (\sin \pi x_1)^2 + (\sin \pi x_2)^2 + (\sin \pi x_3)^2 + (\sin \pi x_4)^2 \text{s.t.} \quad x_i \ge 1 \quad (i = 1 \dots 3) \quad x_4 \ge 3$$

- The NLP is trivially feasible, and
- the value of any feasible solution is non-negative as its objective is the sum of squares.
- In fact, the value of a solution  $(x_1, x_2, x_3, x_4)$  is 0 iff

$$\bullet \ x_1^{x_4} + x_2^{x_4} = x_3^{x_4} \text{, and} \\$$

$$\min (x_1^{x_4} + x_2^{x_4} - x_3^{x_4})^2 + (\sin \pi x_1)^2 + (\sin \pi x_2)^2 + (\sin \pi x_3)^2 + (\sin \pi x_4)^2 \text{s.t.} \quad x_i \ge 1 \quad (i = 1 \dots 3) \quad x_4 \ge 3$$

- The NLP is trivially feasible, and
- the value of any feasible solution is non-negative as its objective is the sum of squares.
- In fact, the value of a solution  $(x_1, x_2, x_3, x_4)$  is 0 iff

• 
$$x_1^{x_4} + x_2^{x_4} = x_3^{x_4}$$
, and

•  $\sin \pi x_i = 0$ , for all  $i = 1 \dots 3$ .

$$\min (x_1^{x_4} + x_2^{x_4} - x_3^{x_4})^2 + (\sin \pi x_1)^2 + (\sin \pi x_2)^2 + (\sin \pi x_3)^2 + (\sin \pi x_4)^2 \text{s.t.} \quad x_i \ge 1 \quad (i = 1 \dots 3) \quad x_4 \ge 3$$

#### Remark

Fermat's Last Theorem is true iff the NLP has optimal value greater than 0.

$$\begin{array}{ll} \min & (x_1^{x_4} + x_2^{x_4} - x_3^{x_4})^2 \\ & + (\sin \pi \, x_1)^2 + (\sin \pi \, x_2)^2 + (\sin \pi \, x_3)^2 + (\sin \pi \, x_4)^2 \\ \text{s.t.} & x_i \ge 1 \quad (i = 1 \dots 3) \\ & x_4 \ge 3 \end{array}$$

#### Remark

Fermat's Last Theorem is true iff the NLP has optimal value greater than 0.

Note: well known that there is an infinite sequence of feasible solutions whose objective value converges to 0!

$$\min (x_1^{x_4} + x_2^{x_4} - x_3^{x_4})^2 + (\sin \pi x_1)^2 + (\sin \pi x_2)^2 + (\sin \pi x_3)^2 + (\sin \pi x_4)^2 \text{s.t.} \quad x_i \ge 1 \quad (i = 1 \dots 3) \quad x_4 \ge 3$$

#### Remark

Fermat's Last Theorem is true iff the NLP has optimal value greater than 0.

Note: well known that there is an infinite sequence of feasible solutions whose objective value converges to 0!

Proving Fermat's Last Theorem amounts to showing that the value 0 can not be attained!

#### Recap

• Non-linear programs are of the form

min 
$$f(x)$$
  
s.t.  $g_1(x) \le 0$   
 $g_2(x) \le 0$   
...  
 $g_m(x) \le 0$ 

where  $f, g_1, \ldots, g_m$  are non-linear functions.

#### Recap

• Non-linear programs are of the form

min f(x)s.t.  $g_1(x) \le 0$   $g_2(x) \le 0$ ...  $g_m(x) \le 0$ ,

where  $f, g_1, \ldots, g_m$  are non-linear functions.

• Non-linear programs are strictly more general than integer programs, and thus likely difficult to solve.

#### Recap

• Non-linear programs are of the form

min f(x)s.t.  $g_1(x) \le 0$   $g_2(x) \le 0$ ...  $g_m(x) \le 0$ ,

where  $f, g_1, \ldots, g_m$  are non-linear functions.

- Non-linear programs are strictly more general than integer programs, and thus likely difficult to solve.
- Some famous questions in Math can easily be reduced to solving certain NLPs