MAT 3007 – Optimization

The Geometry of Linear Problems

Lecture 04

June 15th

Andre Milzarek

iDDA / CUHK-SZ

Repetition

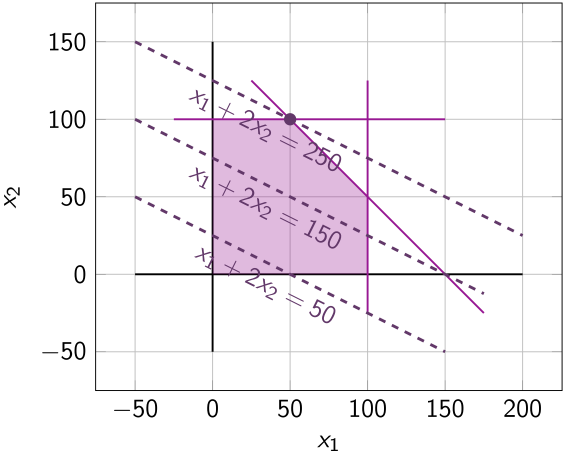
Recap: Linear Problems

- Standard form for LPs.
- We studied several types of problems that can be converted into linear problems.

Geometry of Linear Optimization Problems:

- We solved the production planning problem graphically and observed that the optimal solution appears at the corner of the feasible set.

Recap: Production Planning



- Draw the feasible set and the contour lines of the objective.

$\rightsquigarrow$ Determine the optimal solution.

Polyhedra and Extreme Points

Some Definitions: Polyhedron

Polyhedron

A polyhedron is a set that can be written in the form:

$$\{x \in \mathbb{R}^n : Ax \geq b\},$$

where A is an $m \times n$ matrix and $b \in \mathbb{R}^m$.

- Recall that in the standard form of LP, the feasible set is

$$Ax = b, \quad x \geq 0.$$

- Is this a polyhedron? Why?
- Yes, we can write it as $Ax \geq b$, $Ax \leq b$, $I \cdot x \geq 0$ where I is the identity matrix.

Definition: Convex Set

A set $S \subseteq \mathbb{R}^n$ is convex if for any x, $y \in S$, and any $\lambda \in [0, 1]$, $\lambda x + (1 - \lambda)y \in S$.

Convex Combination

For any $x_1, ..., x_n$ and $\lambda_1, ..., \lambda_n \geq 0$ satisfying $\lambda_1 + \cdots + \lambda_n = 1$, we call $\sum_{i=1}^n \lambda_i x_i$ a convex combination of $x_1, ..., x_n$.

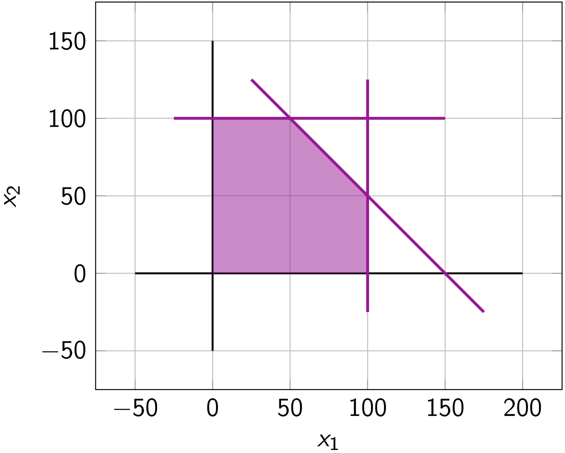
In an LP, the optimal solution tends to be in one of the corners of the feasible region. We first formalize this notion.

Definition: Extreme Point

Let P be a polyhedron. A point $x \in P$ is said to be an extreme point of P if we can not find two vectors $y, z \in P$ with $y, z \neq x$ and a scalar $\lambda \in [0, 1]$, such that $x = \lambda y + (1 - \lambda)z$.

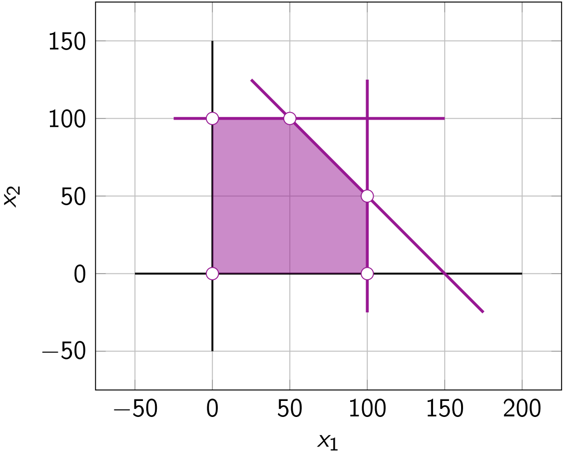
- That is, x cannot be represented as a convex combination of other points in P.
- We sometimes call the extreme point the vertex or corner of the polyhedron.

Example: Extreme Points



How many extreme points are there in this feasible region?

Example: Extreme Points



How many extreme points are there in this feasible region?

- Answer: 5

Extreme Points

We just introduced the definition of extreme points/vertices.

However, this does not tell us how to find those points. We want to have a good way for finding extreme points.

Preview of the Next Steps:

- We need an algebraic way to represent extreme points.
- We will show that it is sufficient to look at extreme points to solve a linear optimization problem.
- Finally, in order find the optimal extreme point, this will lead to the construction of the simplex algorithm for solving LPs.

Finding Extreme Points: Basic Solutions

In the following, we consider an LP in its standard form:

$$\begin{aligned} \text{minimize} \quad & c^\top x \\ \text{subject to} \quad & Ax = b \\ & x \geq 0, \end{aligned}$$

where $x \in \mathbb{R}^n$, A is an $m \times n$ matrix $(m < n)$ and $b \in \mathbb{R}^m$.

General Assumption:

A has linearly independent rows (or equivalently A has full rank m).

What happens if this condition is not satisfied?

- Then either there is a redundant constraints (in which case one can remove it) or the constraints are not consistent (in which case there is no feasible point).

Now, we study the extreme points of an LP in its algebraic form.

Definition: Basic Solution

We call x a basic solution of the LP if and only if

1. $Ax = b$.
2. There exist indices $B(1), ..., B(m)$ such that the columns of

$$\begin{bmatrix} | & | & & | \\ A_{B(1)} & A_{B(2)} & \cdots & A_{B(m)} \\ | & | & & | \end{bmatrix}$$

are linearly independent and $x_i = 0$ for $i \neq B(1), ..., B(m)$.

Procedure to Find a Basic Solution:

1. Choose any m independent columns of A: $A_{B(1)}, ..., A_{B(m)}$.
2. Let $x_i = 0$ for all $i \neq B(1), ..., B(m)$.
3. Solve the equation $Ax = b$ for the remaining $x_{B(1)}, ..., x_{B(m)}$.

Remarks:

- Since $A_{B(1)}, ..., A_{B(m)}$ are linearly independent, the last step must produce a unique solution.
- Basic solution of an LP only depends on its constraints, it has nothing to do with the objective function.

We write

$$A_B = \begin{bmatrix} | & | & & | \\ A_{B(1)} & A_{B(2)} & \cdots & A_{B(m)} \\ | & | & & | \end{bmatrix}, \quad x_B = \begin{bmatrix} x_{B(1)} \\ \vdots \\ x_{B(m)} \end{bmatrix}.$$

Since the columns $A_{B(i)}$ are linearly independent, the matrix A_B is invertible and we have $x_B = A_B^{-1} b$.

- We call $B = \{B(1), ..., B(m)\}$ the basic indices for this basic solution, $A_{B(1)}, ..., A_{B(m)}$ the basic columns, A_B the basis matrix and $x_{B(1)}, ..., x_{B(m)}$ the basic variables.
- We call the remaining indices the non-basic indices, the remaining columns of A the non-basic columns and the remaining variables the non-basic variables.

How many non-zeros could one have in a basic solution (assuming there are m constraints)?

- No more than m!
- Could be anything between 0 to m (typically it is m).

How many basic solutions can one have for a linear program with m constraints and n variables?

- At most $C(n, m) = \frac{n!}{m!(n-m)!}$ (Combination number).
- Therefore for a finite number of linear constraints, there can only be a finite number of basic solutions!

Basic Feasible Solutions

Definition: Basic Feasible Solution

If a basic solution x also satisfies that $x \geq 0$, then we call it a basic feasible solution (BFS).

How to Find a BFS?

- First find a basic solution x.
- Check if $x \geq 0$.

Theorem: Extreme Points and BFS

For the standard LP polyhedron $P := \{x \in \mathbb{R}^n : Ax = b, x \geq 0\}$, the following statements are equivalent:

1. x is an extreme point of P.
2. x is a basic feasible solution.

Recall the production problem:

$$\begin{array}{rrrl} \text{maximize} & x_1 & +2x_2 & \\ \text{subject to} & x_1 & & \leq 100 \\ & & 2x_2 & \leq 200 \\ & x_1 & +x_2 & \leq 150 \\ & x_1, & x_2 & \geq 0 \end{array}$$

with standard form:

$$\begin{array}{rrrrrrl} \text{minimize} & -x_1 & -2x_2 & & & & \\ \text{subject to} & x_1 & & +s_1 & & & = 100 \\ & & 2x_2 & & +s_2 & & = 200 \\ & x_1 & +x_2 & & & +s_3 & = 150 \\ & x_1, & x_2, & s_1, & s_2, & s_3 & \geq 0 \end{array}$$

Example: Continued

We can denote the feasible set by $\{x : Ax = b, x \geq 0\}$, where

$$A = \begin{bmatrix} 1 & 0 & 1 & 0 & 0 \\ 0 & 2 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 \end{bmatrix} \quad b = \begin{bmatrix} 100 \\ 200 \\ 150 \end{bmatrix}.$$

- Choose three independent columns of A, e.g., the first three ($B = \{1, 2, 3\}$), we get the corresponding basic solution via:

$$x_B = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 1 & 0 \end{bmatrix}^{-1} \begin{bmatrix} 100 \\ 200 \\ 150 \end{bmatrix} = \begin{bmatrix} 50 \\ 100 \\ 50 \end{bmatrix}.$$

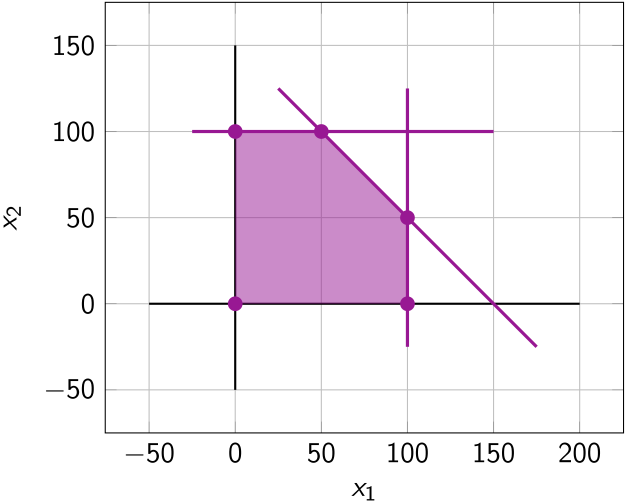
- That is $x_1 = 50, x_2 = 100, s_1 = 50$. Therefore $(50, 100, 50, 0, 0)$ is a basic feasible solution. One can find other basic feasible solutions by choosing other sets of columns.

We can list all basic (feasible) solutions:

Indices	$\{1,2,3\}$	$\{1,2,4\}$	$\{1,2,5\}$
Solution	(50, 100, 50, 0, 0)	(100, 50, 0, 100, 0)	(100, 100, 0, 0, -50)
Status	BFS	BFS	BS, not feasible
Indices	$\{1,3,4\}$	$\{1,4,5\}$	$\{2,3,4\}$
Solution	(150, 0, -50, 200, 0)	(100, 0, 0, 200, 50)	(0, 150, 100, -100, 0)
Status	BS, not feasible	BFS	BS, not feasible
Indices	$\{2,3,5\}$	$\{3,4,5\}$	
Solution	(0, 100, 100, 0, 50)	(0, 0, 100, 200, 150)	
Status	BFS	BFS	

The other two choices $\{1,3,5\}$ and $\{2,4,5\}$ lead to dependent basic columns (therefore no basic solutions can be obtained).

They indeed correspond to all the corners of the feasible set:



Fundamental LP Theorem

Consider a linear problem in standard form and assume that A has full row rank m.

1. If the feasible set is nonempty, there is a basic feasible solution.
2. If there is an optimal solution, there is an optimal solution that is also a basic feasible solution.

$\rightsquigarrow$ In order to find an optimal solution, we only need to look among basic feasible solutions!

Corollary: Characteristics

If an LP with m constraints (in the standard form) has an optimal solution, then there must be an optimal solution with no more than m positive entries.

Consequences:

- We only need to search among basic feasible solutions to find the optimal solution!

How can we search among the basic feasible solutions?

- One may suggest to list all the basic feasible solutions and compare their objective values.

$\rightsquigarrow$ There might be too many BFS!

- For a linear optimization with m constraints and n variables, how many basic feasible solutions can we have?

$\rightsquigarrow$ $C(n, m)$ ... if $n = 1000$, $m = 100$, then this number is 10^{143} ...

The Simplex Method

We need a smarter way to find the optimal solution:

- Simplex method

The simplex method proceeds from one BFS (a corner point of the feasible region) to a neighboring one to continuously improve the value of the objective function until reaching optimality.

- We need to define what it means by adjacent or neighboring solution.
- We need to design an efficient way to find (and move to) the neighboring BFS (e.g., we should try to avoid taking matrix inversions every time).
- We need to design a valid stopping criterion

We will talk the first item and leave the rest to the next lectures.

Neighboring BFS

Definition: Neighboring Basic Solutions

Two basic solutions are neighboring (or adjacent) if they differ by exactly one basic (or non-basic) index.

- For example, a BFS constructed by using the columns $\{1, 2, 3\}$ is a neighbor to the BFS constructed by using the columns $\{1, 3, 5\}$ (but not $\{1, 4, 5\}$).

Consider the production planning problem:

$$\begin{array}{rrrrrrl} \text{minimize} & -x_1 & -2x_2 & & & & \\ \text{subject to} & x_1 & & +s_1 & & & = 100 \\ & & 2x_2 & & +s_2 & & = 200 \\ & x_1 & +x_2 & & & +s_3 & = 150 \\ & x_1, & x_2, & s_1, & s_2, & s_3 & \geq 0. \end{array}$$

The basic (feasible) solutions are given as follows:

Indices	$\{1,2,3\}$	$\{1,2,4\}$	$\{1,2,5\}$
Solution	(50, 100, 50, 0, 0)	(100, 50, 0, 100, 0)	(100, 100, 0, 0, -50)
Objective	-250	-200	infeasible

Indices	$\{1,3,4\}$	$\{1,4,5\}$	$\{2,3,4\}$
Solution	(150, 0, -50, 200, 0)	(100, 0, 0, 200, 50)	(0, 150, 100, -100, 0)
Objective	infeasible	-100	infeasible

Indices	$\{2,3,5\}$	$\{3,4,5\}$	
Solution	(0, 100, 100, 0, 50)	(0, 0, 100, 200, 150)	
Objective	-200	0	

Example

Indices	$\{1,2,3\}$	$\{1,2,4\}$	$\{1,2,5\}$
Solution	(50, 100, 50, 0, 0)	(100, 50, 0, 100, 0)	(100, 100, 0, 0, -50)
Objective	-250	-200	infeasible
Indices	$\{1,3,4\}$	$\{1,4,5\}$	$\{2,3,4\}$
Solution	(150, 0, -50, 200, 0)	(100, 0, 0, 200, 50)	(0, 150, 100, -100, 0)
Objective	infeasible	-100	infeasible
Indices	$\{2,3,5\}$	$\{3,4,5\}$	
Solution	(0, 100, 100, 0, 50)	(0, 0, 100, 200, 150)	
Objective	-200	0	

Say we start from $\{3,4,5\}$ and only allow to move from one BFS to its neighbor (and we move only if the obj. value decreases). Then:

- If we choose $\{1,4,5\}$ first, then we can move along:

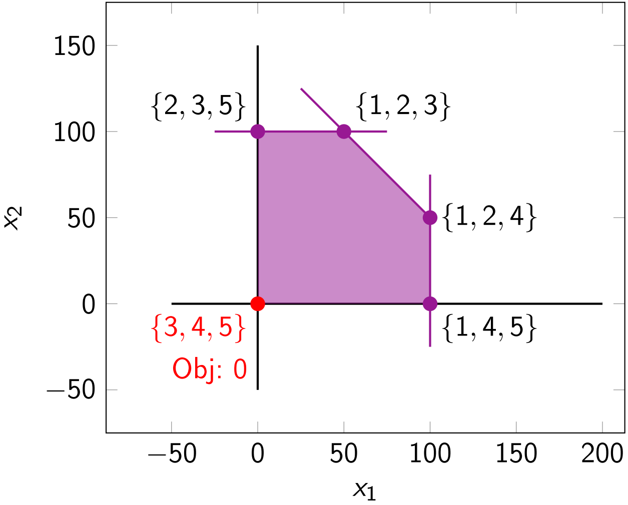
$$\{3,4,5\} \to \{1,4,5\} \to \{1,2,4\} \to \{1,2,3\}$$

- We can also start from $\{2,3,5\}$, which will give us

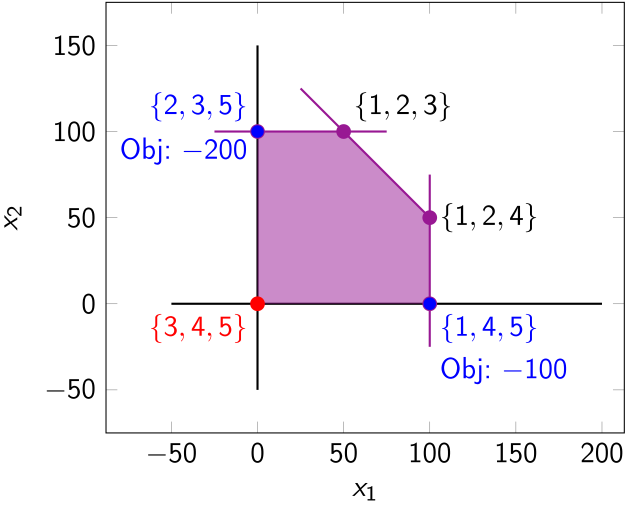
$$\{3,4,5\} \to \{2,3,5\} \to \{1,2,3\}$$

- Both get to the optimal solution in finite steps.

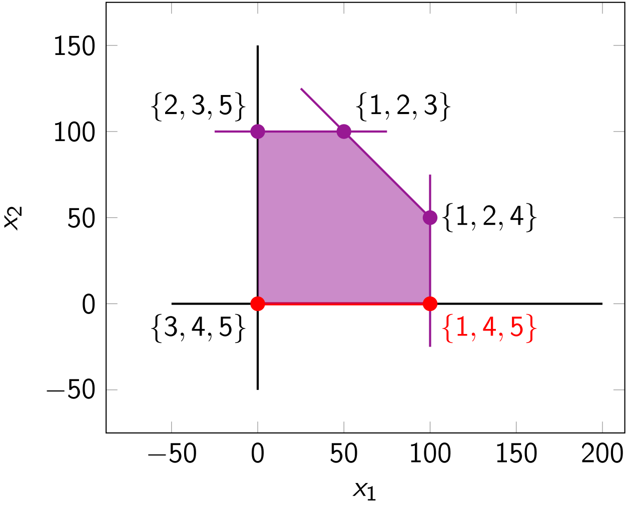
Example



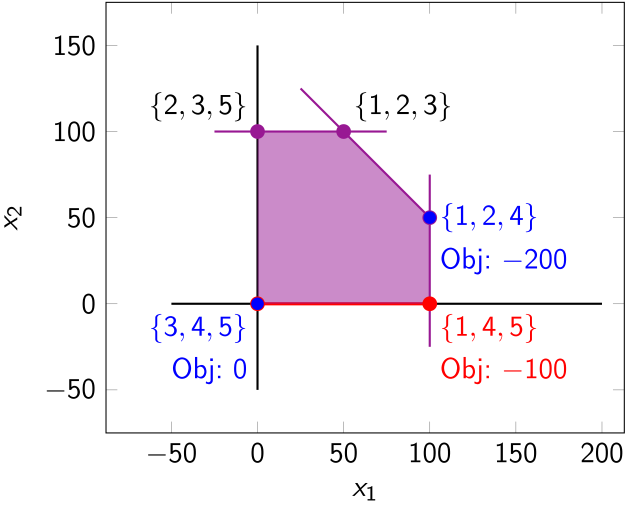
Example


150
100
50
0
−50
−50
0
50
100
150
200
x_1
x_2
{2, 3, 5}
Obj: −200
{1, 2, 3}
{1, 2, 4}
{3, 4, 5}
{1, 4, 5}
Obj: −100

Example

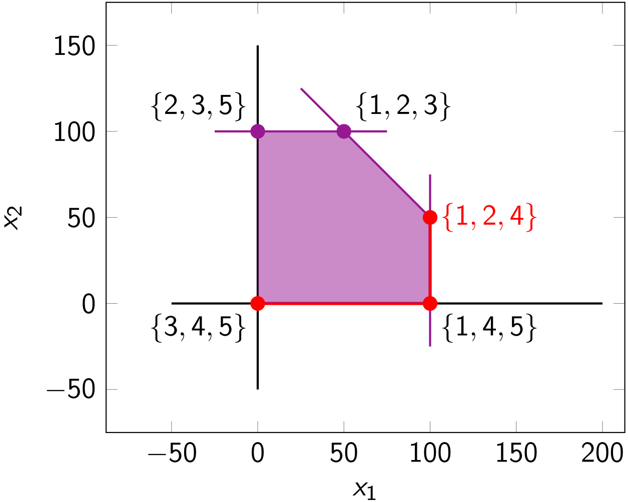


Example

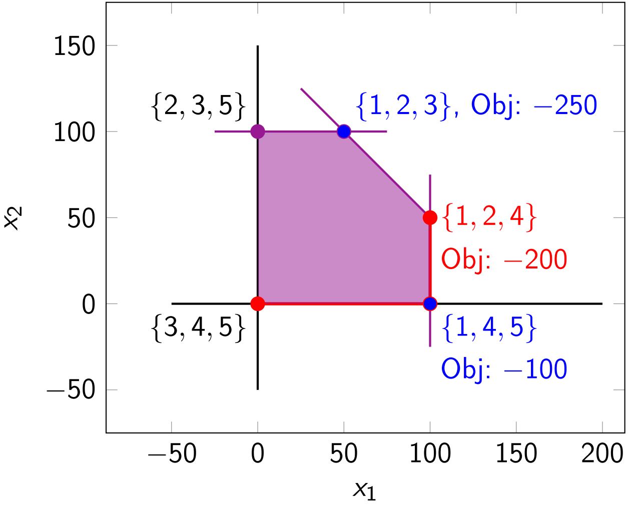


150
100
50
0
−50
−50
0
50
100
150
200
x_1
x_2
$\{2, 3, 5\}$
$\{1, 2, 3\}$
$\{1, 2, 4\}$
Obj: −200
$\{3, 4, 5\}$
Obj: 0
$\{1, 4, 5\}$
Obj: −100

Example

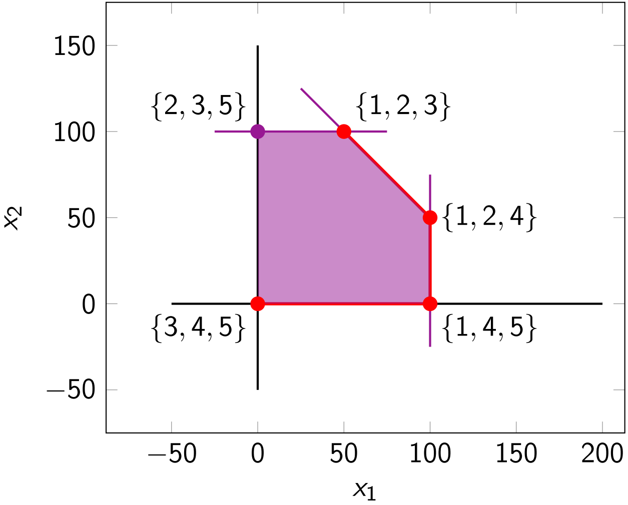


Example



150
100
50
0
−50
x_2
$\{2, 3, 5\}$
$\{1, 2, 3\}$, Obj: -250
$\{1, 2, 4\}$
Obj: -200
$\{3, 4, 5\}$
$\{1, 4, 5\}$
Obj: -100
−50
0
50
100
150
200
x_1

Example



The main question is how to find a neighboring BFS that improves the current one (reduces the objective function):

- We only need to show how to do it for one step. Then, we can simply iterate the same methods (until we reach optimal).

A naive way is simply to check all neighbors of the current BFS:

- There are $m(n-m)$ potential neighbors for a BFS (choosing one index to leave, choosing one to enter). For each one, we need to solve a linear equation of m variables.
- This works, but is likely to be very slow.

We want a more efficient way to do it $\rightsquigarrow$ next lectures!

Questions?