

Solution to CO327 (2021Spring) Assignment 3

Lecturer: Andersen Ang

July 2, 2021

1 Quadratic program of MAXCUT (14 points)

Consider the following undirected graph.

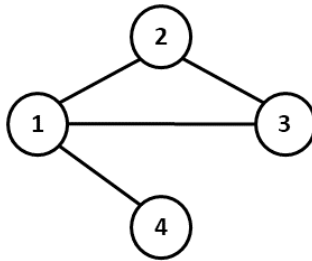


Figure 1: A undirected graph.

A quick review of graph theory: a graph G is a pair of sets V, E , where V is a set whose elements are called vertices, and E is a set of paired vertices, whose elements are called edges. For the graph in Fig.1, $V = \{1, 2, 3, 4\}$ and $E = \{(1, 2), (1, 3), (1, 4), (2, 1), (2, 3), (3, 2)\}$ or simplified as $E = \{(1, 2), (1, 3), (1, 4), (2, 3)\}$.

1. Write down the 0-1 adjacency matrix \mathbf{A} of this graph.
2. Consider the MAXCUT problem on this graph, write down the capacity for the problem in the following forms

$$\max_{\mathbf{x}} \underbrace{\frac{1}{2} \sum_{i,j} A_{ij}(1 - x_i x_j)}_{c_1(\mathbf{x})} \quad \text{s.t. } x_i \in \{-1, +1\}, \quad \text{and} \quad \max_{\mathbf{x}} \underbrace{\frac{1}{4} \sum_{i,j} A_{ij}(x_i - x_j)^2}_{c_2(\mathbf{x})} \quad \text{s.t. } x_i \in \{-1, +1\}.$$

that is, write down the cost functions $c_1(\mathbf{x})$ and $c_2(\mathbf{x})$.

3. The capacity can also be expressed as $\mathbf{x}^\top \mathbf{Q} \mathbf{x}$, find \mathbf{Q} .
4. Given a solution with max capacity.

Sol

$$\mathbf{A} = \begin{bmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}, \quad (4 \text{ pt: } 1 \text{ points for each correct column or row})$$

The functions c

$$\begin{aligned}
2c_1(\mathbf{x}) &= (1 - x_1x_2) + (1 - x_1x_3) + (1 - x_1x_4) \\
&+ (1 - x_2x_1) + (1 - x_2x_3) \\
&+ (1 - x_3x_1) + (1 - x_3x_2) \\
&+ (1 - x_4x_1)
\end{aligned}
\quad , \quad
\begin{aligned}
4c_2(\mathbf{x}) &= (x_1 - x_2)^2 + (x_1 - x_3)^2 + (x_1 - x_4)^2 \\
&+ (x_2 - x_1)^2 + (x_2 - x_3)^2 \\
&+ (x_3 - x_1)^2 + (x_3 - x_2)^2 \\
&+ (x_4 - x_1)^2
\end{aligned}$$

4 points each (total 8 points): 1/2 point for each bracket

$$\mathbf{Q} = \begin{bmatrix} 4 & 0 & 0 & 0 & 0 \\ 0 & -0.5 & -0.5 & -0.5 & 0 \\ 0 & -0.5 & 0 & -0.5 & 0 \\ 0 & -0.5 & -0.5 & 0 & 0 \\ 0 & -0.5 & 0 & 0 & 0 \end{bmatrix} \text{ for } \mathbf{x}^\top \mathbf{Q} \mathbf{x} \text{ or } \frac{-1}{2} \begin{bmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \end{bmatrix} \text{ for } 4 - \mathbf{x}^\top \mathbf{Q} \mathbf{x}$$

this question has error so everybody get 4 pts

Solution with max capacity Multiple sol. with max capacity 3: $\{1\}, \{1, 2\}, \{1, 3\}$.

2 points: 1 for capacity and 1 for any sol. If give more than 1 sol, give +1 bonus point

2 Linear location planning (19 points)

A country somewhere in Europe is planning on where to build hospital(s). By abstracting the geographical information, the map of the country can be represented by the following weighted undirected graph in Fig.2.

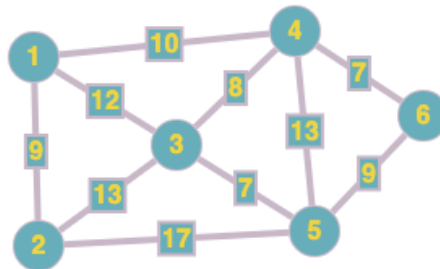


Figure 2: A weighted undirected graph.

For the nodes $i \in V = \{1, 2, \dots, 6\}$, each node represents a city in this country. The number on the edge $(i, j) \in E$ represents the distance between the two connected cities.

1. Write down E (the set of edges) and the 0-1 adjacency matrix \mathbf{A} of the vertices of this graph.
2. Now, each city can only build at most one hospital. All the hospitals have infinite capacity, and are able to serve (i) all the patients of the city it is located; and (ii) all the patients from its neighbouring cities, if the distance to the neighbouring city is less than or equal to 10.

Determine the minimum number of hospital to build such that each city is covered by a hospital. Write this down as an optimization problem. *You do not need to solve the problem.

3. Now, the road between city 1 and city 3 has been improved, the number 12 is changed to 2.

Now, **on top of the conditions stated in Q2**, the country is deciding where to build crematoriums (place to burning dead bodies). Again, each city can only build at most one crematorium. All the crematoriums have infinite capacity, and are able to serve (i) the city it is located; and (ii) all its neighbouring cities (i.e., they are connected by an edge in the graph).

However, due to religious reasons:

- People do not want to have a crematorium if a hospital is located there. That is, if there is an hospital located in the city i , there should not be a crematorium.
- People do not want to have a crematoriums being “too close” to each other. That is, if the **shortest** distance between two cities is less than or equal to 16, then there can be at most 1 crematorium in either one of these two cities.

Determine the minimum number of hospitals and crematoriums to build such that each city is covered by a hospital and a crematorium. Write this down as an optimization problem. *You do not need to solve the problem.

Sol

1. $E = \{(1, 2), (1, 3), (1, 4), (2, 3), (2, 5), (3, 4), (3, 5), (4, 5), (4, 6), (5, 6)\}$ (1 points if all correct)

2

$$\mathbf{A}_{\text{adj}} = \begin{bmatrix} 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 \end{bmatrix} \quad (3 \text{ points, } -0.5 \text{ points for each wrong row/col})$$

$$\begin{array}{ll} \min_{\mathbf{x}} & \mathbf{1}^\top \mathbf{x} \\ \text{s.t.} & \mathbf{x} \in \{0, 1\}^6 \\ & \mathbf{A}\mathbf{x} \geq \mathbf{1} \end{array} \quad \mathbf{A} = \begin{bmatrix} 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 \end{bmatrix} \quad (7 \text{ points})$$

(1 point for min, 1 point for cost function, 1 point for integer constraint, 1 point for inequality constant, 3 points for \mathbf{A} (-0.5 points for each wrong row/col))

3

$$\begin{array}{ll} \min_{\mathbf{x}, \mathbf{y}} & \mathbf{1}^\top \mathbf{x} + \mathbf{1}^\top \mathbf{y} \\ \text{s.t.} & \mathbf{x}, \mathbf{y} \in \{0, 1\}^6 \\ & \mathbf{A}\mathbf{x} \geq \mathbf{1} \\ & \mathbf{A}_{\text{adj}}\mathbf{y} \geq \mathbf{1} \\ & \mathbf{B}\mathbf{y} \leq \mathbf{1} \end{array} \quad \mathbf{A} = \begin{bmatrix} 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 \end{bmatrix} \quad \mathbf{B} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 & 1 \end{bmatrix} \quad (8 \text{ points})$$

(1 point for cost function, 1 point for integer constraint, 1 point for each inequality constant, 1 point for each matrix (-0.5 points for each wrong row/col))

3 Nonlinear location planning (11 points)

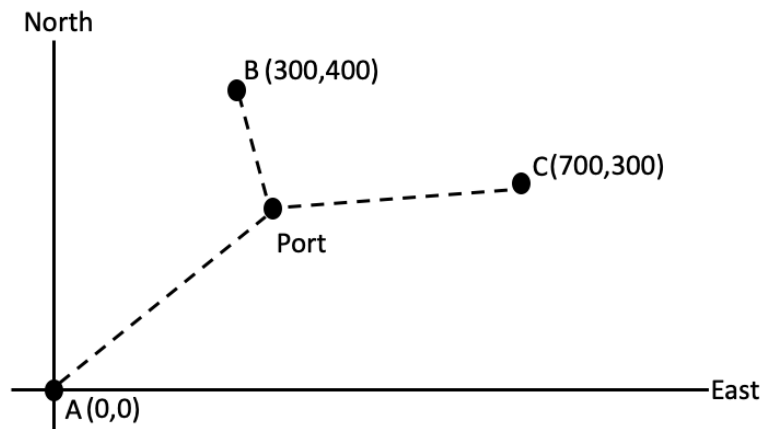
Your oil company wants to build a refinery port that will be supplied from 3 port cities A,B and C:

- Port B is located 300km east and 400km north of Port A,
- Port C is located 400km east and 100km south of Port B,

Your task is to determine the location of the refinery port so that the total amount of pipe required to connect the refinery port to the ports is minimized.

1. Consider a two-dimensional Cartesian coordinate system (x-y plane). Suppose Port A is located at the coordinate (0,0). Draw the locations of Port B and Port C.
2. Suppose the location of the refinery is located at the unknown coordinate (x_1, x_2) , write down the total amount of Euclidean distance between the refinery to the ports. (Hint: the Euclidean distance between two points $P(x_1, x_2)$ and $Q(y_1, y_2)$ is $\sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2}$).
3. Write down the optimization problem.
4. In this course we do not focus how to solve nonlinear programming problems. Describe how would you approach this problem (explain what you will do on trying to solve this problem).

Sol



1. The map: 3 points, 1 point each for each dot (A,B,C)
2. Total distance: 3 points, 1 point for each square-root term

$$\sqrt{x_1^2 + x_2^2} + \sqrt{(x_1 - 300)^2 + (x_2 - 400)^2} + \sqrt{(x_1 - 700)^2 + (x_2 - 300)^2}$$

3. The optimization problem: 3 point or 0 point, see below

$$\min_{x_1, x_2, x_3} \sqrt{x_1^2 + x_2^2} + \sqrt{(x_1 - 300)^2 + (x_2 - 400)^2} + \sqrt{(x_1 - 700)^2 + (x_2 - 300)^2}$$

Important remarks

- there is NO constraint for this problem. For example, if x_2 is negative, it only means the port is located in the south of A. So if the solution has any extra constraint, like $x_1 \geq 0$, no point awarded for this part.
- Since LP with no constraint is unbounded, so students may think this problem is unbounded. However, no constraints doesn't mean the problem is unbounded for nonlinear problems.

4

- Any theoretically correct solution or explanation scores 2 points. Correct solution includes:
 - Find the Fermat - Torricelli point
 - Use iterative method like gradient descent method / Newton's iteration
- Solution that makes sense but may be theoretically incorrect / practically inefficient : 1 point. These include any center of triangle other than the Fermat - Torricelli point: e.g. centroid, center of mass, center of gravity, incenter, etc.
- Nonsense scores 0 point.

4 Constrained location planning: Radio placement (40 points)

You are a radio engineer and your task is to determine where to place two large radio towers (with sizes $a_1 \times b_1$ and $a_2 \times b_2$, respectively). Refer to Fig.3, due to some urban constraints, each radio tower can only be placed within a certain region. For the blue tower, it can not be placed outside the box with the corners $(0, 0)$, $(0, l_{22})$, (l_{12}, l_{22}) and $(l_{12}, 0)$. Similarly, the red tower not be placed outside the box with the corners (l_{13}, l_{23}) , (l_{13}, l_{24}) , (l_{14}, l_{24}) and (l_{14}, l_{23}) .

To avoid interference of the signal, you want to build the towers as far away as possible. Your job now is to determine where to place these towers such that the “distance” between them is maximized, while making sure that the towers are located within their feasible building ranges.

1. Suppose now we are measuring the distance between the towers using Euclidean distance, write down the optimization problem. (Hint: the Euclidean distance between two points $P(x_1, x_2)$ and $Q(y_1, y_2)$ is $\sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2}$).
2. Suppose now we are measuring the distance between the towers using the so-called “min distance”, write down the optimization problem. (Hint: the “min distance” between two points $P(x_1, x_2)$ and $Q(y_1, y_2)$ is defined as $\min\{x_1 - y_1, x_2 - y_2\}$).
3. Convert the problem in part 2 into a linear program.

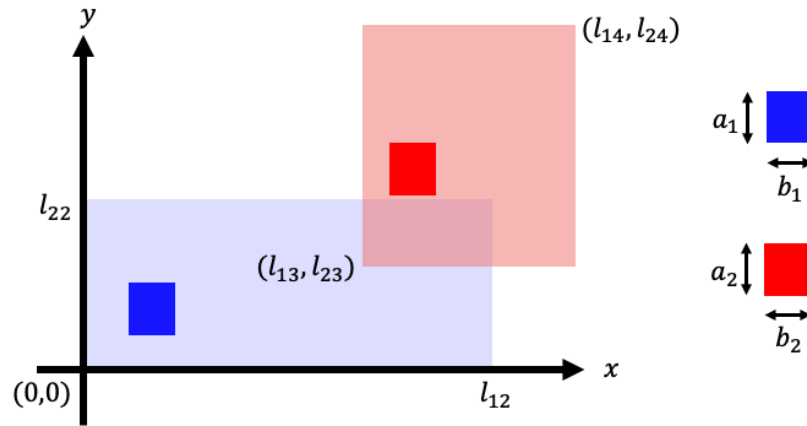


Figure 3: Layout of the feasible region for the radio towers.

4. Suppose now we are now measuring the distance between the towers using the so-called “taxicab distance”, write down the optimization problem. (Hint: the “taxicab distance” between two points $P(x_1, x_2)$ and $Q(y_1, y_2)$ is defined as $|x_1 - y_1| + |x_2 - y_2|$).
5. Convert the problem in part 4 into a linear program. (Hint: you need to introduce new variables.)

Sol

Important note: I use the notation $\mathbf{x} = [x_1, x_2]$, $\mathbf{y} = [y_1, y_2]$ to represent the points, NOT $[x_1, y_1]$, $[x_2, y_2]$, so be-careful of notation mistake.

- 1 Let (x_1, x_2) be the lower left corner of the block box, and let (y_1, y_2) be the lower left corner of the block box. (2 points for decision variables)

$$\begin{aligned} \max_{x_1, y_1, x_2, y_2} \quad & \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2} \quad \text{s.t.} \quad x_1 \geq 0, y_1 \geq 0 \\ & x_1 + a_1 \leq l_{12}, \quad x_2 + b_1 \leq l_{22} \\ & y_1 \geq l_{13}, \quad y_2 \geq l_{23} \\ & y_1 + a_2 \geq l_{14}, \quad y_2 + b_2 \geq l_{24} \end{aligned} \quad (6 \text{ points})$$

1 for max, 1 for cost function, 4 points for constraints (1 point for each coordinate)

2

$$\max_{x_1, y_1, x_2, y_2} \min\{x_1 - y_1, x_2 - y_2\} \quad \text{s.t.} \quad \text{same constraints in 1} \quad (7 \text{ points})$$

1 for max, 2 for cost function, 4 for constraints

3 Let $z = \max_{x_1, y_1, x_2, y_2} z$ (1 point)

$$\begin{aligned} \max_{x_1, y_1, x_2, y_2, z} \quad & z \quad \text{s.t.} \quad x_1 - y_1 \geq z \\ & x_2 - y_2 \geq z \\ & \text{plus the constraints in 1} \end{aligned} \quad (8 \text{ points})$$

(1 for max 1 for cost function, 2+4 for constraints)

Note If the student define the point as the upper right corner of the box, similar marking applies as long as the constraints and the modelling are correct.

4

$$\max_{x_1, x_2, y_1, y_2} |x_1 - y_1| + |x_2 - y_2| \quad \text{s.t. same constraints in 1} \quad (6 \text{ points})$$

1 for max, 1 for cost function, 4 for constraints

5 Let $z \leq |x_1 - y_1|$ and $w \leq |x_2 - y_2|$. (2 points)

$$\begin{aligned} \max_{w, z} \quad & w + z \quad \text{s.t.} \quad x_1 - y_1 \geq z \\ & -(x_1 - y_1) \leq -z \\ & x_2 - y_2 \geq w \\ & -(x_2 - y_2) \leq -w \\ & \text{plus the constraints in 1 on } x_i, y_i \end{aligned} \quad (8 \text{ points})$$

1 for max, 1 for cost function, 8 for constraints

5 The bodyguards of president (10 points)

You are the president of the United States. Being such a VIP (Very important person), you are always followed by 3 bodyguards with the following specifications: for each bodyguard,

- For security, the maximum distance between you and the guard cannot exceed 5 meters.
- For convenience, you want to have at least 1 meter of distance between you and the guard.

Now, your task is to arrange the location of the three bodyguards, subject to the specification above. In addition, to prevent forming a “blind spot” (all the guards stay in the same spot), you order the guards to be away from each other as further as possible.

1. Formulate this problem as an optimization problem.
2. Draw the feasible set.
3. Guess what is the geometrical shape of such arrangement, explain why such geometrical shape is optimal.

Note: you only need to consider the problem in a two-dimensional space.

Sol

In a coordinate system, assume you (the VIP) are located at $(0, 0)$, and let $(x_i, y_i), i \in \{1, 2, 3\}$ be the location of the bodyguards (1 point for the decision variable).

$$\max_{\mathbf{x}, \mathbf{y}} \sum_{i=1}^3 \sqrt{x_i^2 + y_i^2} \quad \text{s.t.} \quad 1 \leq \sqrt{x_i^2 + y_i^2} \leq 5 \quad \forall i. \quad (4 \text{ points})$$

(1 point for max, 1 point for correct cost function, 2 points for constraint)

Note: if the students do not put the VIP at location $(0, 0)$ then they need to introduce extra coordinate. Then as long as their expression is correct they get the points.

- Shape: equilateral triangle (1 point).
- For any explanation making sense (1 points).
- If the explanation is sounds or theoretically correct (bonus for 1 points).

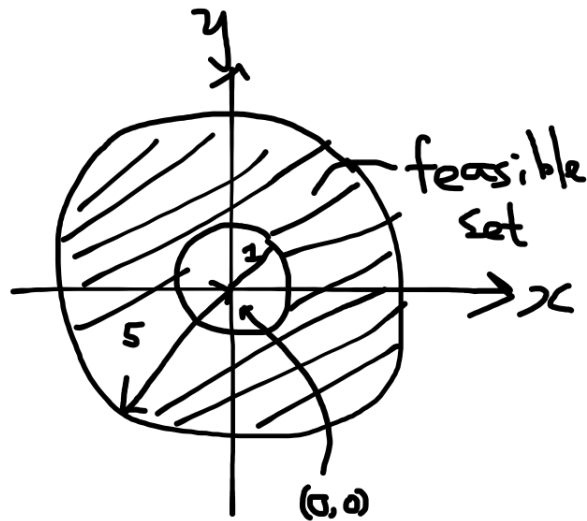


Figure 4: Total 3 points: 1 point for the annulus shape, 2 points for the radius.

6 Auction market (17 points)

Consider the following betting on horse racing.

Order	Price limit	Quantity limit	Horse 1	Horse 2	Horse 3	Horse 4	Horse 5
1	0.75	10	1	1	1		
2	0.35	5				1	1
3	0.4	10	1		1		1
4	0.95	10	1	1	1	1	
5	0.75	5		1		1	

1. Formulate it as the Pari-mutuel model in the original form
2. Formulate the model in 1 to a linear program

Sol

Decision variable : $\mathbf{x} \in \mathbb{R}^5$ (1 point).

$$\max_{\mathbf{x}} \left\{ \boldsymbol{\pi}^\top \mathbf{x} - \max_j \mathbf{a}_{:j}^\top \mathbf{x} \right\} \text{ s.t. } \mathbf{0} \leq \mathbf{x} \leq \mathbf{q} \quad (5 \text{ points})$$

$$(\max 1 \text{ pt.}, \boldsymbol{\pi}^\top \mathbf{x} 1 \text{ pt.}, -\max_j \mathbf{a}_{:j}^\top \mathbf{x} 1 \text{ pt.}, \mathbf{0} \leq \mathbf{x} 1 \text{ pt.}, \mathbf{x} \leq \mathbf{q} 1 \text{ pt.})$$

where

$$\boldsymbol{\pi} = \begin{bmatrix} 0.75 \\ 0.35 \\ 0.4 \\ 0.95 \\ 0.75 \end{bmatrix}, \quad \mathbf{q} = \begin{bmatrix} 10 \\ 5 \\ 10 \\ 10 \\ 5 \end{bmatrix}, \quad \mathbf{A} = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 \end{bmatrix}, \quad (3 \text{ points}, 1 \text{ point each})$$

LP form: introduce a new variable $y = \max_j \mathbf{a}_j^\top \mathbf{x}$ (1 point)

$$\max_{\mathbf{x}, y} \boldsymbol{\pi}^\top \mathbf{x} - y \text{ s.t. } \mathbf{0} \leq \mathbf{x} \leq \mathbf{q}, \mathbf{A}^\top \mathbf{x} \leq \mathbf{1}y \quad (5 \text{ points})$$

(max 1 pt., $\boldsymbol{\pi}^\top \mathbf{x}$ 1 pt., $\boldsymbol{\pi}^\top \mathbf{x} - y$ 1 pt., $\mathbf{0} \leq \mathbf{x} \leq \mathbf{q}$ 1 pt., $\mathbf{A}^\top \mathbf{x} \leq \mathbf{1}y$ 1 pt.)
 where \mathbf{A} is the same as last part (1 pt.) and $\mathbf{1}$ is a all-one vector in \mathbb{R}^5 . (1 pt.)

7 Formulation of image matching (24 points)

Introduction In this question, you are going to formulate a quadratic integer program for an engineering problem called “Image Matching” (IM).

In IM, you are given two images, your task is to find the “correspondences” between these two images. A way to find such correspondences is to model the images using graphs. Refer to the example shown in Fig.5, two input images are represented as two graphs respectively. Let $G = (V, E), G' = (V', E')$ be the graphs representing these two images, where the symbol G denotes the graph, the symbol V represents the set of the vertices in the graph, and E represents the set of edges in the graph, and the apostrophe $'$ indicating that we are referring to the second image.

Note: In this question, you do not need to worry about how G and G' are obtained, you assume someone working in a tech company have already prepared G and G' for you, i.e., you are given G and G' in the beginning. Furthermore, you are given a set of values s_V and s_E , to be discussed later.

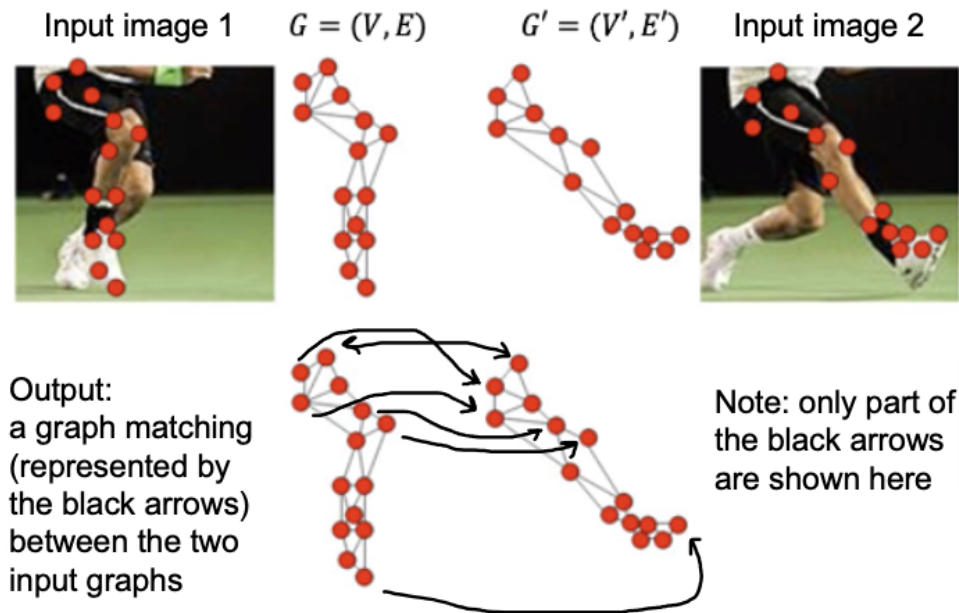


Figure 5: An illustrative example of image matching. Top row: two input images and their graph representation. Bottom row: the output of solving a QIP, which is the matching between the vertices of the two graphs. The image is modified from an image from a research paper.

On the decision variable Given G and G' , your job is to determine the correspondence between V and V' . That is, for the n nodes in V and n' nodes in V' , your task is to determine which node in V' is connected to the nodes in V .

1. Based on the above description, what is the decision variable for this problem?

The goal of IM is to determine the correspondence between V and V' . One way to determine such correspondence is to model it by forming an optimization problem on maximizing a cost function subject to some constraints.

On the cost function To model the cost function for the IM between two graphs, a common approach is to use a “similarity score” that measures the similarity between the two graphs. As a graph has two parts (the vertices and the edge), such similarity score has two parts:

- The first part represents the similarity between two vertices: one vertex from V and one vertex from V' . Mathematically, for a vertex-pair (v_i, v_a) , where $v_i \in V$ and $v_a \in V'$, the score function can be represented as c_{ia} .
- The second part represents the similarity between two edges: one edge from E and one edge from E' . Mathematically, for a edge-pair (e_{ij}, e_{ab}) , where $e_{ij} \in E$ and $e_{ab} \in E'$, the score function can be represented as $d_{ia,jb}$.

You do not need to worry about how the value in c and d are obtained, you assume someone working in a tech company have already prepared the values for you.

2. Based on the above description, write down the cost function.

On the constraint Now we talk about the constraint. There is only a simple constraint in IM, the matching has to be “one-to-one”, that is, for all v in V , it has to be matched to exactly one v' in V' , and for all v' in V' , it has to be matched to exactly one v in V

3. Based on the above description, write down the constraints.

The optimization problem **4. Write down the whole optimization problem.**

On the cost function: matrix form representation The similarity functions are usually represented by a symmetric matrix \mathbf{A} , where a non-diagonal element $A_{ia,jb} = d_{ia,jb}$ contains the edge similarity of two correspondences (v_i, v'_a) and (v_j, v'_b) , and a diagonal term $\mathbf{A}_{ia,ia} = c_{ia}$ represents the node similarity of a correspondence (v_i, v_a) .

5. Based on the above description, rewrite down the cost function and write down the optimization problem again with the new cost function.

Sol

Note. This question might be difficult and therefore it is highly possible that students cannot get it all correct. For each part, the points are given in multiple of 4 that:

- If they got it fully correct, they get all the points.
- If they do not give the exact correct expression, but give a similar explanation or expression to the correct solution, they get 75% of the points.

- If they do not give the exact correct expression, but with more-or-less correct explanation, they get 50% of the points.
- If they do not give the correct solution, but they tried, they get 25% of the points.
- Nonsense or no solution get no point.

The solution:

1. 4 points for a matrix $\mathbf{x} \in \{0, 1\}^{n \times n'}$
2. 8 points for $\sum_{ia} c_{ia} x_{ia} + \sum_{iabj} d_{ia,jb} x_{ia} x_{bj}$ (4 points for each term)
3. 4 points for $\sum_i x_{ia} = 1 \forall a$ and 4 points for $\sum_a x_{ia} = 1 \forall i$. If $=$ is replaced by \leq , it is also considered to be correct.
4. The optimization problem = combine parts 1,2,3. The problem is a max problem (4 points for the max).
5. 4 points for $\max \sum A_{iajb} x_{ia} x_{jb}$

END.