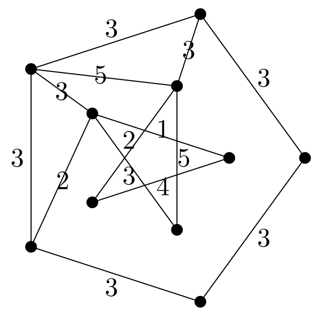
1. (Enumeration (25 points)) A DNA sequence is a word on the letters $\{A, T, G, C\}$. Let a_n be the number of DNA sequences of length n such that the letter 'T' is not followed by the letter 'G'.
   (a) Write a recurrence relation for a_n.
   (b) Solve the recurrence relation and give an explicit formula for a_n. What is the number of these sequences of length 10.

Solution:

(a) For each sequence of length $n-1$ which ends in a letter different from T, one obtains four sequences of length n with no T followed by G. From all the sequences of length $n-1$ which end in T one can add only three symbols at the end. There are a_{n-2} sequences of length $n-1$ which end in T. Hence
$$a_n = 4(a_{n-1} - a_{n-2}) + 3a_{n-2} = 4a_{n-1} - a_{n-2}, n \geq 3.$$
We clearly have $a_1 = 4$ and $a_2 = 15$ (from the 16 sequences of length two only the sequence TG is forbidden).

(b) The characteristic polynomial of the recurrence is
$$x^2 - 4x + 1,$$
which has roots $x_1 = 2 + \sqrt{3}$ and $x_2 = 2 - \sqrt{3}$. The general solution is
$$a_n = \alpha(2 + \sqrt{3})^n + \beta(2 - \sqrt{3})^n.$$
In order to determine the values of α and β we use the initial values:
$$4 = \alpha(2 + \sqrt{3}) + \beta(2 - \sqrt{3})$$
$$15 = \alpha(2 + \sqrt{3})^2 + \beta(2 - \sqrt{3})^2.$$
By multiplying the first equation by $(2 - \sqrt{3})$ and substracting it from the second one we obtain $\alpha = \frac{7+4\sqrt{3}}{2\sqrt{3}(2+\sqrt{3})}$ and $\beta = \frac{4-\alpha(2+\sqrt{3})}{2-\sqrt{3}} = -\frac{7-4\sqrt{3}}{2\sqrt{3}(2-\sqrt{3})}$. Therefore
$$a_n = \frac{12 + 7\sqrt{3}}{6}(2 + \sqrt{3})^{n-1} + \frac{12 - 7\sqrt{3}}{6}(2 - \sqrt{3})^{n-1}.$$
In particular
$$a_{10} = 564719.$$

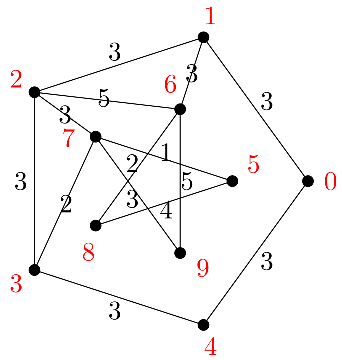
2. (Graphs (25 points)
   (a) What is a spanning tree? Given a graph G, is there a unique spanning tree? Explain your answer.
   (b) Run the Kruskal algorithm on the weighted graph G given below. Write down the individual steps and the output.
   (c) Will the output of the Kruskal algorithm be unique when it is run on graphs with distinct positive edge weights? Explain why.



Solution:

(a) A spanning tree is a tree T which is a subgraph of G and $V(T) = V(G)$.
If G is not connected then clearly G has no spanning trees (a tree is connected).
A graph is connected if and only if it has a spanning tree. If G is not connected there are no spanning trees in G. If G is connected and it is a tree itself then it has an only spanning tree. Otherwise, it has a cycle and, from any given spanning tree $T \subset G$ one can build a different spanning tree by adding one edge from $E(G) \setminus E(T)$ (which will create a cycle) and removing one edge of T in this cycle: what remains is still connected and has $n - 1$ edges, so it is a tree. Therefore G has more than one spanning tree.

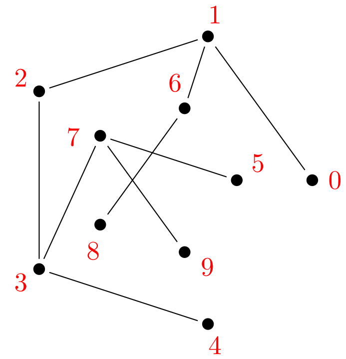
(b) Label the vertices as in the Figure,



and sort the edges in nondecreasing order (the ordering in this case is not unique). The output of the algorithm is described in the following table, where at each step we add the next edge in the ordering as long as it produces an acyclic graph:

i	e_i	weight	$E(X_i)$
1	{5,7}	1	{5,7}
2	{6,8}	2	{5,7}, {6,8}
3	{3,7}	2	{5,7}, {6,8},{3,7}
4	{0,1}	3	{5,7}, {6,8},{3,7},{0,1}
5	{1,2}	3	{5,7}, {6,8},{3,7},{0,1},{1,2}
6	{2,3}	3	{5,7}, {6,8},{3,7},{0,1},{1,2},{2,3}
7	{3,4}	3	{5,7}, {6,8},{3,7},{0,1},{1,2},{2,3},{3,4}
8	{4,0}	3	{5,7}, {6,8},{3,7},{0,1},{1,2},{2,3},{3,4}
9	{2,7}	3	{5,7}, {6,8},{3,7},{0,1},{1,2},{2,3},{3,4}
10	{7,9}	3	{5,7}, {6,8},{3,7},{0,1},{1,2},{2,3},{3,4},{7,9}
11	{1,6}	3	{5,7}, {6,8},{3,7},{0,1},{1,2},{2,3},{3,4},{7,9} {1,6}
12	{5,8}	4	{5,7}, {6,8},{3,7},{0,1},{1,2},{2,3},{3,4},{7,9} {1,6}
13	{2,6}	5	{5,7}, {6,8},{3,7},{0,1},{1,2},{2,3},{3,4},{7,9} {1,6}
14	{3,9}	5	{5,7}, {6,8},{3,7},{0,1},{1,2},{2,3},{3,4},{7,9} {1,6}

The output is the spanning tree containing the edges listed at the last step:



with total weight 23.

(c) If the edges have pairwise distinct weights then the minimum spanning tree is unique: suppose that T_1, T_2 are two distinct trees with miniimum weight; suppose that T_1 contains the edge e with smallest weight in $E(T_1) \cup E(T_2)$; add this edge to T_2 and remove one edge of T_2 forming a cycle distinct from e; we obtain a spanning tree with smaller weight than T_2 (and T_1), a contradiction.

3. (Linear programing (25 points)) We must run a program P and we have two available machines M_1 and M_2. Each execution of P spends $4s$ on M_1 and $1s$ on M_2. The cost per execution on M_1 is 1 cents and in M_2 is 4 cent. The energy consumption per execution on M_1 is $3mW$ and on M_2 it is $1mW$. We must run at least 10 executions in total but we can not exceed $37s$ of running time nor 28 cents of cost. We want to minimize the energy consumption.

   (a) Write a Linear Program to solve the problem.
   (b) Draw the feasible region of the problem and identify a solution graphically.
   (c) Write the Linear Program in equational form and run the simplex method on it.

Solution:

(a) Let x and y be the number of executions of P on M_1 and M_2 respectively. The Linear Program is

Minimize:	$3x + y$	Energy consumption
Subject to:	$4x + y \leq 37$	Time
	$x + 4y \leq 28$	Cost
	$x + y \geq 10$	Number of executions
	$x \geq 0, y \geq 0$	Nonnegativity

(b) The feasible region is indicated in Figure 1, the intersection of the semiplanes defined by the restrictions. The objective function is represented as a dashed line. Its minimum is at the intersection of a parallel line to the dashed line which first touches the feasinle region. Therefore its minimum occurs at the point $A = (4, 6)$ in the figure. The minimum is $18mW$.

(c) In the equational form we should maximize a linear function subject to constraints that are all equalities except for the variables, which must be nonnegative. We introduce slack variables

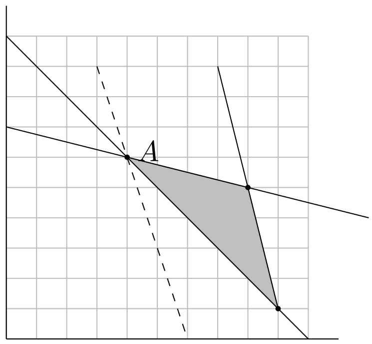


Figure 1: The feasible region and the objective function.

z, w, t and write the problem as

$$\begin{aligned} \text{Maximize:} \quad & -3x - y \\ \text{Subject to:} \quad & 4x + y + z = 37 \\ & x + 4y + w = 28 \\ & -x - y + t = -10 \\ & x, y, z, w, t \geq 0 \end{aligned}$$

The matrix of the system is

$$A = \begin{pmatrix} 4 & 1 & 1 & 0 & 0 \\ 1 & 4 & 0 & 1 & 0 \\ -1 & -1 & 0 & 0 & 1 \end{pmatrix}$$

In order to find a basic feasible solution to start the simplex algorithm, we select two variables and set them to 0 as long as the corresponding solution to the system has nonnegative entries. The choice $w = t = 0$ gives the basic feasible solution $(4, 6, 15, 0, 0)$ and the tableau

$$\begin{array}{llll|l} x & = 4 & -(1/3)w & +(4/3)t & \\ y & = 6 & -(1/3)w & -(1/3)t & \\ z & = 15 & -w & -5t & \\ \hline c & = -18 & -(2/3)w & -(11/3)t & -18 \end{array}$$

We can not increase the objective function by replacing the values $w = t = 0$ by any positive values (because the coefficients of c are negative), so that the given basic feasible condition is already the optimal one.

4. (Non-linear Optimisation 2. (25 points))

   Let $f(x, y) = (x^2 + y^2) + ((x-4)^2 + (y-2)^2) + ((x-1)^2 + (y-4)^2)$, which gives the sum of the squares of the distances from a point (x, y) in the plane to the points $\mathbf{a} = (0, 0)$, $\mathbf{b} = (4, 2)$ and $\mathbf{c} = (1, 4)$.

   (a) Show that $f(x, y)$ is a strictly convex function on $\mathbb{R}^2$.

   (b) Find the point $\mathbf{d} = (d_1, d_2)$ which minimizes the sum of the squares of the distances to $\mathbf{a}, \mathbf{b}$ and $\mathbf{c}$.

   (c) Starting at the point $(1, 1)$, find the point in the first iteration of the gradient method for the function f.

Solution:

(a) The function $g(x, y) = (x - x_0)^2 + (y - y_0)^2$ is convex in $\mathbb{R}^2$ for each fixed $(x_0, y_0) \in \mathbb{R}^2$. The function f is the sum of three such functions and it is therefore convex in $\mathbb{R}^2$. To see that it is strictly convex, we compute the Hessian of f and show that it is positive definite.

$$\begin{aligned}\nabla f(x, y) &= (2x + 2(x - 4) + 2(x - 1), 2y + 2(y - 2) + 2(y - 4) \\ &= (6x - 10, 6y - 12).\end{aligned}$$

and

$$Hf(x, y) = \begin{pmatrix} 6 & 0 \\ 0 & 6 \end{pmatrix}.$$

The principal determinants of H are $6, 36$, both positive (or the eigenvalues are both positive). Hence, Hf is positive definite.

(b) Since f is convex in $\mathbb{R}^2$, it has a global minimum, which must be a critical point of f. From the above computation of the gradient, we have

$$\begin{aligned}6x - 10 &= 0 \\ 6y - 12 &= 0\end{aligned}$$

which gives the point $d = (5/3, 2)$.

(c) By starting at the point $\mathbf{x}_0 = (1, 1)$ the first point in the iteration of the gradient method gives

$$\mathbf{x}_1 = \mathbf{x}_0 + t\nabla f(\mathbf{x}_0) = (1, 1) - t(-4, -6),$$

where t is the point which minimizes the function

$$g(t) = f(\mathbf{x}_0 - t\nabla f(\mathbf{x}_0)) = f(1 + 4t, 1 + 6t) = 156t^2 - 52t + 21,$$

which is the solution of $g'(t) = 312t - 52 = 0$, namely $t = 1/6$. Therefore the first point in the iteration of the gradient method is

$$\mathbf{x}_1 = (1, 1) - (1/6)(-4, -6) = (5/3, 2).$$