Optimization

G. V. V. Sharma

Contents

1 Definition

1.1. Express the problem of finding the distance of the point $\mathbf{P} = \begin{pmatrix} 3 \\ -5 \end{pmatrix}$ from the line

$$L: \quad \begin{pmatrix} 3 & -4 \end{pmatrix} \mathbf{x} = 26 \tag{1.1.1}$$

as an optimization problem.
Solution: The given problem can be expressed as

$$\min_{\mathbf{x}} g(\mathbf{x}) = \|\mathbf{x} - \mathbf{P}\|^2 \tag{1.1.2}$$

$$\text{s.t.} \quad \mathbf{n}^T \mathbf{x} = c \tag{1.1.3}$$

where

$$\mathbf{n} = \begin{pmatrix} 3 \\ -4 \end{pmatrix} \tag{1.1.4}$$

$$c = 26 \tag{1.1.5}$$

1.2. Explain Problem 1.1 through a plot and find a graphical solution.

1.3. Solve (1.1.2) using cvxpy.
Solution: The following code yields

```
codes/opt/line_dist_cvx.py
```

$$\mathbf{x}_{\min} = \begin{pmatrix} 2.64 \\ -4.52 \end{pmatrix}, \tag{1.3.1}$$

$$g(\mathbf{x}_{\min}) = 0.6 \tag{1.3.2}$$

1.4. Convert (1.1.2) to an *unconstrained* optimization problem.
Solution: L in (1.1.1) can be expressed in terms of the direction vector $\mathbf{m}$ as

$$\mathbf{x} = \mathbf{A} + \lambda \mathbf{m}, \tag{1.4.1}$$

where $\mathbf{A}$ is any point on the line and

$$\mathbf{m}^T \mathbf{n} = 0 \tag{1.4.2}$$

Substituting (1.4.1) in (1.1.2), an unconstrained optimization problem

$$\min_{\lambda} f(\lambda) = \|\mathbf{A} + \lambda \mathbf{m} - \mathbf{P}\|^2 \tag{1.4.3}$$

is obtained.

1.5. Solve (1.4.3).
Solution:

$$f(\lambda) = (\lambda \mathbf{m} + \mathbf{A} - \mathbf{P})^T (\lambda \mathbf{m} + \mathbf{A} - \mathbf{P}) \tag{1.5.1}$$

$$= \lambda^2 \|\mathbf{m}\|^2 + 2\lambda \mathbf{m}^T (\mathbf{A} - \mathbf{P}) + \|\mathbf{A} - \mathbf{P}\|^2 \tag{1.5.2}$$

$$\because f^{(2)} \lambda = 2 \|\mathbf{m}\|^2 > 0 \tag{1.5.3}$$

the minimum value of $f(\lambda)$ is obtained when

$$f^{(1)}(\lambda) = 2\lambda \|\mathbf{m}\|^2 + 2\mathbf{m}^T (\mathbf{A} - \mathbf{P}) = 0 \tag{1.5.4}$$

$$\implies \lambda_{\min} = -\frac{\mathbf{m}^T (\mathbf{A} - \mathbf{P})}{\|\mathbf{m}\|^2} \tag{1.5.5}$$

Choosing $\mathbf{A}$ such that

$$\mathbf{m}^T (\mathbf{A} - \mathbf{P}) = 0, \tag{1.5.6}$$

substituting in (1.5.5),

$$\lambda_{\min} = 0 \quad \text{and} \tag{1.5.7}$$
$$\mathbf{A} - \mathbf{P} = \mu\mathbf{n} \tag{1.5.8}$$

for some constant μ. (1.5.8) is a consequence of (1.4.2) and (1.5.6). Also, from (1.5.8),

$$\mathbf{n}^T(\mathbf{A}-\mathbf{P}) = \mu\left\|\mathbf{n}\right\|^2 \tag{1.5.9}$$
$$\implies \mu = \frac{\mathbf{n}^T\mathbf{A} - \mathbf{n}^T\mathbf{P}}{\left\|\mathbf{n}\right\|^2} = \frac{c - \mathbf{n}^T\mathbf{P}}{\left\|\mathbf{n}\right\|^2} \tag{1.5.10}$$

from (1.1.3). Substituting $\lambda_{\min} = 0$ in (1.4.3),

$$\min_{\lambda} f(\lambda) = \left\|\mathbf{A}-\mathbf{P}\right\|^2 = \mu^2\left\|\mathbf{n}\right\|^2 \tag{1.5.11}$$

upon substituting from (1.5.8). The distance between $\mathbf{P}$ and L is then obtained from (1.5.11) as

$$\left\|\mathbf{A}-\mathbf{P}\right\| = |\mu|\left\|\mathbf{n}\right\| \tag{1.5.12}$$
$$= \frac{\left|\mathbf{n}^T\mathbf{P} - c\right|}{\left\|\mathbf{n}\right\|} \tag{1.5.13}$$

after substituting for μ from (1.5.10). Using the corresponding values from Problem (1.1) in (1.5.13),

$$\min_{\lambda} f(\lambda) = 0.6 \tag{1.5.14}$$

2 Convex Function

2.1. The following python script plots

$$f(\lambda) = a\lambda^2 + b\lambda + d \tag{2.1.1}$$

for

$$a = \left\|\mathbf{m}\right\|^2 > 0 \tag{2.1.2}$$
$$b = \mathbf{m}^T(\mathbf{A}-\mathbf{P}) \tag{2.1.3}$$
$$c = \left\|\mathbf{A}-\mathbf{P}\right\|^2 \tag{2.1.4}$$

where $\mathbf{A}$ is the intercept of the line L in (1.1.1) on the x-axis and the points

$$\mathbf{U} = \begin{pmatrix}\lambda_1\\ f(\lambda_1)\end{pmatrix}, \mathbf{V} = \begin{pmatrix}\lambda_2\\ f(\lambda_2)\end{pmatrix} \tag{2.1.5}$$
$$\mathbf{X} = \begin{pmatrix}t\lambda_1 + (1-t)\lambda_2\\ f\left[t\lambda_1 + (1-t)\lambda_2\right]\end{pmatrix}, \tag{2.1.6}$$
$$\mathbf{Y} = \begin{pmatrix}t\lambda_1 + (1-t)\lambda_2\\ tf(\lambda_1) + (1-t)f(\lambda_2)\end{pmatrix} \tag{2.1.7}$$

for

$$\lambda_1 = -3, \lambda_2 = 4, t = 0.3 \tag{2.1.8}$$

in Fig. 2.1. Geometrically, this means that any point $\mathbf{Y}$ between the points $\mathbf{U}, \mathbf{V}$ on the line UV is always above the point $\mathbf{X}$ on the curve $f(\lambda)$. Such a function f is defined to be *convex* function

```
codes/opt/1.2.py
```

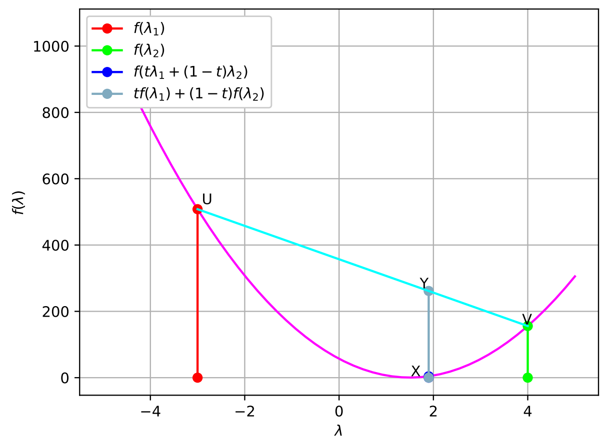

Fig. 2.1. $f(\lambda)$ versus λ

.

2.2. Show that

$$f\left[t\lambda_1 + (1-t)\lambda_2\right] \le tf(\lambda_1) + (1-t)f(\lambda_2) \tag{2.2.1}$$

for $0 < t < 1$. This is true for any convex function.

2.3. Show that

$$(2.2.1) \implies f^{(2)}(\lambda) > 0 \tag{2.3.1}$$

2.4. Show that a covex function has a unique minimum.

3 Gradient Descent

3.1. Find a numerical solution for (2.1.1)
Solution: A numerical solution for (2.1.1) is obtained as

$$\lambda_{n+1} = \lambda_n - \mu f'(\lambda_n) \tag{3.1.1}$$
$$= \lambda_n - \mu(2a\lambda_n + b) \tag{3.1.2}$$

where λ_0 is an inital guess and μ is a variable parameter. The choice of these parameters is very important since they decide how fast the algorithm converges.

3.2. Write a program to implement (3.1.2).
Solution: Download and execute

```
codes/opt/gd.py
```

3.3. Find a closed form solution for λ_n in (3.1.2) using the one sided Z transform.

3.4. Find the condition for which (3.1.2) converges, i.e.

$$\lim_{n\to\infty} |\lambda_{n+1} - \lambda_n| = 0 \tag{3.4.1}$$

4 Lagrange Multipliers

4.1. Find

$$\min_{\mathbf{x}} g(\mathbf{x}) = \|\mathbf{x} - \mathbf{P}\|^2 = r^2 \tag{4.1.1}$$

$$\text{s.t.} \quad h(\mathbf{x}) = \mathbf{n}^T \mathbf{x} - c = 0 \tag{4.1.2}$$

by plotting the circles $g(\mathbf{x})$ for different values of r along with the line $g(\mathbf{x})$.
Solution: The following code plots Fig. 4.1

```
codes/opt/concirc.py
```

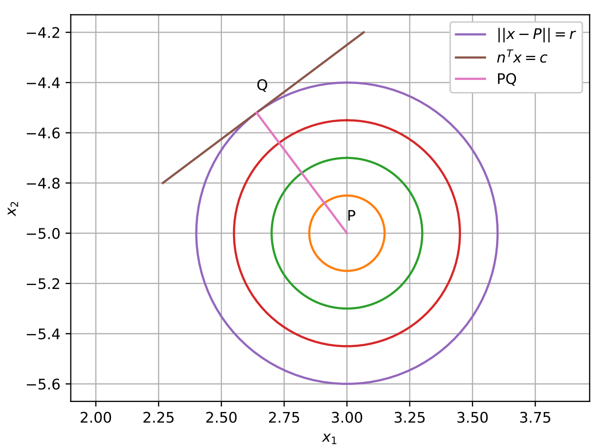

Fig. 4.1. Finding $\min_{\mathbf{x}} g(\mathbf{x})$

.

4.2. By solving the quadratic equation obtained from (4.1.1), show that

$$\min_{\mathbf{x}} r = \frac{3}{5}, \mathbf{x}_{\min} = \mathbf{Q} = \begin{pmatrix} 2.64 \\ -4.52 \end{pmatrix} \tag{4.2.1}$$

In Fig. 4.1, it can be seen that $\mathbf{Q}$ is the point of contact of the line L with the circle of minimum radius $r = \frac{3}{5}$.

4.3. Show that

$$\nabla h(\mathbf{x}) = \begin{pmatrix} 3 \\ -4 \end{pmatrix} = \mathbf{n} \tag{4.3.1}$$

where

$$\nabla = \begin{pmatrix} \frac{\partial}{\partial x_1} \\ \frac{\partial}{\partial x_2} \end{pmatrix} \tag{4.3.2}$$

4.4. Show that

$$\nabla g(\mathbf{x}) = 2\left\{\mathbf{x} - \begin{pmatrix} 3 \\ -5 \end{pmatrix}\right\} = 2\{\mathbf{x} - \mathbf{P}\} \tag{4.4.1}$$

4.5. From Fig. 4.1, show that

$$\nabla g(\mathbf{Q}) = \lambda \nabla h(\mathbf{Q}), \tag{4.5.1}$$

Solution: In Fig. 4.1, PQ is the normal to the line L, represented by $h(\mathbf{x})$. $\because$ the normal vector of L is in the same direction as PQ, for some constant k,

$$(\mathbf{Q} - \mathbf{P}) = k\mathbf{n} \tag{4.5.2}$$

which is the same as (4.5.1) after substituting from (4.3.1). and (4.4.1).

4.6. Use (4.5.1) and $\mathbf{h}(\mathbf{Q}) = 0$ from (4.1.2) to obtain $\mathbf{Q}$.
Solution: From the given equations, we obtain

$$(\mathbf{Q} - \mathbf{P}) - \lambda\mathbf{n} = 0 \tag{4.6.1}$$

$$\mathbf{n}^T \mathbf{Q} - c = 0 \tag{4.6.2}$$

which can be simplifed to obtain

$$\begin{pmatrix} \mathbf{I} & -\mathbf{n} \\ \mathbf{n}^T & 0 \end{pmatrix} \begin{pmatrix} \mathbf{Q} \\ \lambda \end{pmatrix} = \begin{pmatrix} \mathbf{P} \\ c \end{pmatrix} \tag{4.6.3}$$

The following code computes the solution to (4.6.3)

```
codes/opt/lagmul.py
```

4.7. Define

$$C(\mathbf{x}, \lambda) = g(\mathbf{x}) - \lambda h(\mathbf{x}) \tag{4.7.1}$$

and show that $\mathbf{Q}$ can also be obtained by solving the equations

$$\nabla C(\mathbf{x}, \lambda) = 0. \tag{4.7.2}$$

What is the sign of λ? C is known as the Lagrangian and the above technique is known as the Method of Lagrange Multipliers.

4.8. Obtain $\mathbf{Q}$ using gradient descent.
Solution:

```
codes/opt/gd_lagrange.py
```

5 KKT Conditions

5.1 Equality Constraints

5.1.1 Find

$$\min_{\mathbf{x}} f(\mathbf{x}) = \left\| \mathbf{x} - \begin{pmatrix} 8 \\ 6 \end{pmatrix} \right\|^2 = r^2 \tag{5.1.1.1}$$

$$\text{s.t.} \quad g(\mathbf{x}) = \begin{pmatrix} 1 & 1 \end{pmatrix} \mathbf{x} - 9 = 0 \tag{5.1.1.2}$$

by plotting the circles $f(\mathbf{x})$ for different values of r along with the line $g(\mathbf{x})$.

Solution: The following code plots Fig. 5.1.1

```
codes/kkt/2.1.py
```

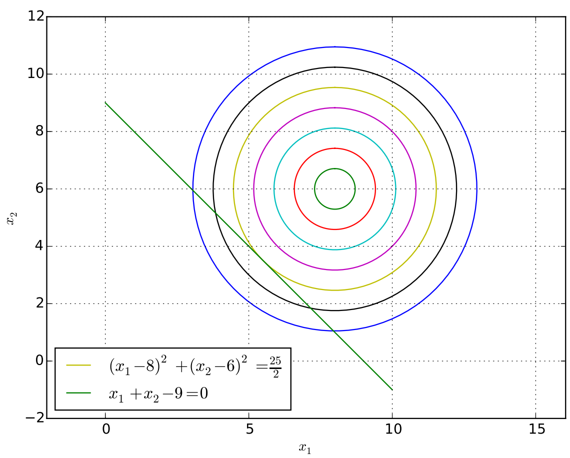

Fig. 5.1.1. Finding $\min_{\mathbf{x}} f(\mathbf{x})$

.

5.1.2 Show that

$$\min r = \frac{5}{\sqrt{2}} \tag{5.1.2.1}$$

5.1.3 Show that

$$\nabla g(\mathbf{x}) = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \tag{5.1.3.1}$$

where

$$\nabla = \begin{pmatrix} \frac{\partial}{\partial x_1} \\ \frac{\partial}{\partial x_2} \end{pmatrix} \tag{5.1.3.2}$$

5.1.4 Show that

$$\nabla f(\mathbf{x}) = 2 \left\{ \mathbf{x} - \begin{pmatrix} 8 \\ 6 \end{pmatrix} \right\} \tag{5.1.4.1}$$

5.1.5 From Fig. 5.1.1, show that

$$\nabla f(\mathbf{p}) = \lambda \nabla g(\mathbf{p}), \tag{5.1.5.1}$$

where $\mathbf{p}$ is the point of contact.

5.1.6 Use (5.1.5.1) and $\mathbf{g}(\mathbf{p}) = 0$ from (5.1.1.2) to obtain $\mathbf{p}$.

5.1.7 Define

$$L(\mathbf{x}, \lambda) = f(\mathbf{x}) - \lambda g(\mathbf{x}) \tag{5.1.7.1}$$

and show that $\mathbf{p}$ can also be obtained by solving the equations

$$\nabla L(\mathbf{x}, \lambda) = 0. \tag{5.1.7.2}$$

What is the sign of λ? L is known as the Lagrangian and the above technique is known as the Method of Lagrange Multipliers.

Solution:

```
codes/kkt/2.3.py
```

5.2 Inequality Constraints

5.2.1 Modify the code in problem 5.1.1 to find a graphical solution for minimising

$$f(\mathbf{x}) \tag{5.2.1.1}$$

with constraint

$$g(\mathbf{x}) \geq 0 \tag{5.2.1.2}$$

Solution: This problem reduces to finding the radius of the smallest circle in the shaded area in Fig. 5.2.1 . It is clear that this radius is 0.

```
codes/kkt/2.4.py
```

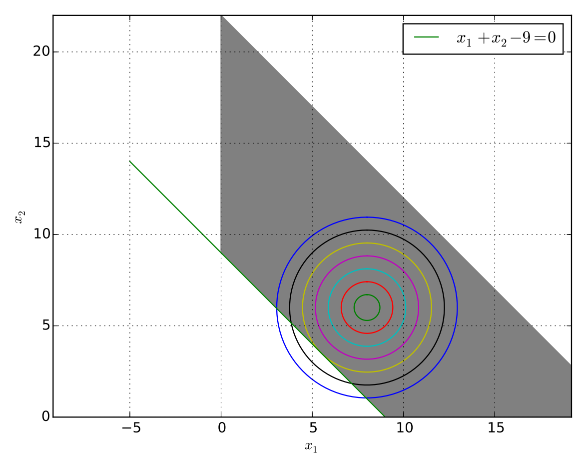

Fig. 5.2.1. Smallest circle in the shaded region is a point.

5.2.2 Now use the method of Lagrange multipliers to solve problem 5.2.1 and compare with the graphical solution. Comment.

Solution: Using the method of Lagrange multipliers, the solution is the same as the one obtained in problem 5.2.1, which is different from the graphical solution. This means that the Lagrange multipliers method cannot be applied blindly.

5.2.3 Repeat problem 5.2.2 by keeping $\lambda = 0$. Comment.

Solution: Keeping $\lambda = 0$ results in $\mathbf{x} = \begin{pmatrix} 8 \\ 6 \end{pmatrix}$, which is the correct solution. The minimum value of $f(\mathbf{x})$ without any constraints lies in the region $g(\mathbf{x}) = 0$. In this case, $\lambda = 0$.

5.2.4 Find a graphical solution for minimising

$$f(\mathbf{x}) \tag{5.2.4.1}$$

with constraint

$$g(\mathbf{x}) \leq 0 \tag{5.2.4.2}$$

Summarize your observations.

Solution: In Fig. 5.2.4, the shaded region represents the constraint. Thus, the solution is the same as the one in problem 5.2.1. This implies that the method of Lagrange multipliers can be used to solve the optimization problem with this inequality constraint as well. Table 5.2.4 summarizes the conditions for this based on the observations so far.

```
codes/kkt/2.7.py
```

Cost	Constraint	λ
$f(\mathbf{x})$	$g(\mathbf{x}) = 0$	< 0
	$g(\mathbf{x}) \geq 0$	0
	$g(\mathbf{x}) \leq 0$	< 0

TABLE 5.2.4
SUMMARY OF CONDITIONS.

5.2.5 Find a graphical solution for

$$\min_{\mathbf{x}} f(\mathbf{x}) = \left\| \mathbf{x} - \begin{pmatrix} 8 \\ 6 \end{pmatrix} \right\|^2 \tag{5.2.5.1}$$

with constraint

$$g(\mathbf{x}) = \begin{pmatrix} 1 & 1 \end{pmatrix} \mathbf{x} - 18 = 0 \tag{5.2.5.2}$$

Solution:

```
codes/kkt/2.8.py
```

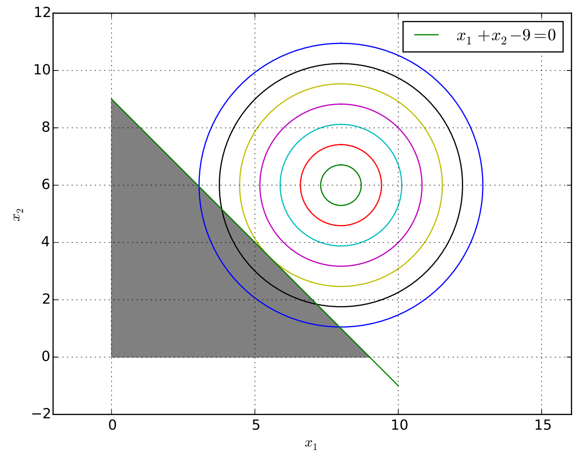

Fig. 5.2.4. Finding $\min_{\mathbf{x}} f(\mathbf{x})$.

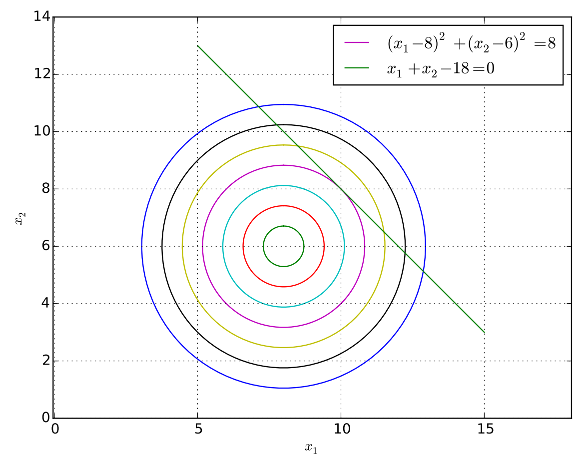

Fig. 5.2.5. Finding $\min_{\mathbf{x}} f(\mathbf{x})$.

5.2.6 Repeat problem 5.2.5 using the method of Lagrange mutipliers. What is the sign of λ?

Solution: Using the following python script, λ is positive and the minimum value of f is 8.

```
codes/kkt/2.9.py
```

5.2.7 Solve

$$\min_{\mathbf{x}} f(\mathbf{x}) \tag{5.2.7.1}$$

with constraint

$$g(\mathbf{x}) \geq 0 \tag{5.2.7.2}$$

Solution: Since the unconstrained solution is outside the region $g(\mathbf{x}) \geq 0$, the solution is the same as the one in problem 5.2.5.

5.2.8 Based on the problems so far, generalise the Lagrange multipliers method for

$$\min_{\mathbf{x}} f(\mathbf{x}), \quad g(\mathbf{x}) \geq 0 \tag{5.2.8.1}$$

Solution: Considering $L(\mathbf{x}, \lambda) = f(\mathbf{x}) - \lambda g(\mathbf{x})$, for $g(\mathbf{x}) = \begin{pmatrix} 1 & 1 \end{pmatrix}\mathbf{x} - 18 \geq 0$ we found $\lambda > 0$ and for $g(\mathbf{x}) = \begin{pmatrix} 1 & 1 \end{pmatrix}\mathbf{x} - 9 \leq 0, \lambda < 0$. A single condition can be obtained by framing the optimization problem as

$$\min_{\mathbf{x}} f(\mathbf{x}), \quad g(\mathbf{x}) \leq 0 \tag{5.2.8.2}$$

with the Lagrangian

$$L(\mathbf{x}, \lambda) = f(\mathbf{x}) + \lambda g(\mathbf{x}), \tag{5.2.8.3}$$

provided

$$\nabla L(\mathbf{x}, \lambda) = 0 \Rightarrow \lambda > 0 \tag{5.2.8.4}$$

else, $\lambda = 0$.

5.3 Karush Kuhn-Tucker Conditions

5.3.1 Solve

$$\min_{\mathbf{x}} f(\mathbf{x}) = \mathbf{x}^T \begin{pmatrix} 4 & 0 \\ 0 & 2 \end{pmatrix} \mathbf{x} \tag{5.3.1.1}$$

with constraints

$$g_1(\mathbf{x}) = \begin{pmatrix} 3 & 1 \end{pmatrix}\mathbf{x} - 8 = 0 \tag{5.3.1.2}$$

$$g_2(\mathbf{x}) = 15 - \begin{pmatrix} 2 & 4 \end{pmatrix}\mathbf{x} \geq 0 \tag{5.3.1.3}$$

Solution: Considering the Lagrangian

$$\nabla L(\mathbf{x}, \lambda, \mu) = 0 \tag{5.3.1.4}$$

resulting in the matrix equation

$$\Rightarrow \begin{pmatrix} 8 & 0 & 3 & 2 \\ 0 & 4 & 1 & 4 \\ 3 & 1 & 0 & 0 \\ 2 & 4 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \lambda \\ \mu \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 8 \\ 15 \end{pmatrix} \tag{5.3.1.5}$$

$$\Rightarrow \begin{pmatrix} x_1 \\ x_2 \\ \lambda \\ \mu \end{pmatrix} = \begin{pmatrix} 1.7 \\ 2.9 \\ -3.12 \\ -2.12 \end{pmatrix} \tag{5.3.1.6}$$

using the following python script. The (incorrect) graphical solution is available in Fig. 5.3.1

```
codes/kkt/2.12.py
```

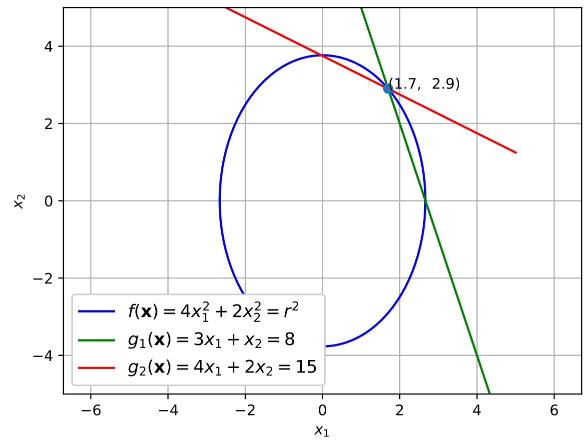

Fig. 5.3.1. Incorrect solution is at intersection of all curves $r = 5.33$

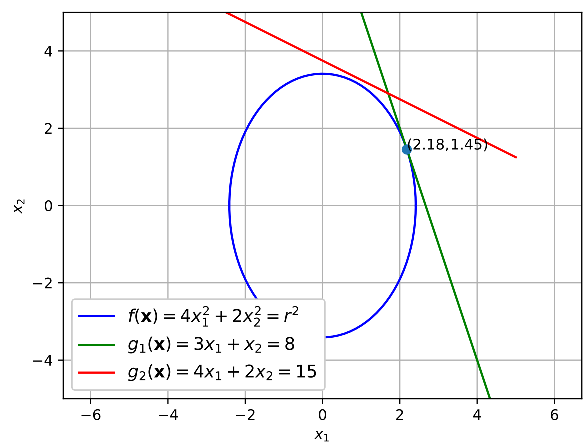

Fig. 5.3.2. Optimal solution is where $g_1(x)$ touches the curve $r = 4.82$

Note that $\mu < 0$, contradicting the necessary condition in (5.2.8.4).

5.3.2 Obtain the correct solution to the previous problem by considering $\mu = 0$.

5.3.3 Solve

$$\min_{\mathbf{x}} f(\mathbf{x}) \tag{5.3.3.1}$$

with constraints

$$g_1(\mathbf{x}) = 0 \tag{5.3.3.2}$$

$$g_2(\mathbf{x}) \leq 0 \tag{5.3.3.3}$$

5.3.4 Based on whatever you have done so far, list the steps that you would use in general for solving a convex optimization problem like (5.3.1.1) using Lagrange Multipliers. These are called Karush-Kuhn-Tucker(KKT) conditions.

Solution: For a problem defined by

$$\mathbf{x}^* = \min_{\mathbf{x}} f(\mathbf{x}) \tag{5.3.4.1}$$

$$\text{subject to } h_i(\mathbf{x}) = 0, \forall i = 1, .., m \tag{5.3.4.2}$$

$$\text{subject to } g_i(\mathbf{x}) \leq 0, \forall i = 1, .., n \tag{5.3.4.3}$$

the optimal solution is obtained through

$$\mathbf{x}^* = \min_{\mathbf{x}} L(\mathbf{x}, \lambda, \mu) \tag{5.3.4.4}$$

$$= \min_{\mathbf{x}} f(\mathbf{x}) + \sum_{i=1}^{m} \lambda_i h_i(\mathbf{x}) + \sum_{i=1}^{n} \mu_i g_i(\mathbf{x}), \tag{5.3.4.5}$$

using the KKT conditions

$$\Rightarrow \nabla_{\mathbf{x}} f(\mathbf{x}) + \sum_{i=1}^{m} \nabla_{\mathbf{x}} \lambda_i h_i(\mathbf{x}) + \sum_{i=1}^{n} \mu_i \nabla_{\mathbf{x}} g_i(\mathbf{x}) = 0 \tag{5.3.4.6}$$

$$\text{subject to } \mu_i g_i(\mathbf{x}) = 0, \forall i = 1, .., n \tag{5.3.4.7}$$

$$\text{and } \mu_i \geq 0, \forall i = 1, .., n \tag{5.3.4.8}$$

5.3.5 Maxmimize

$$f(\mathbf{x}) = \sqrt{x_1 x_2} \tag{5.3.5.1}$$

with the constraints

$$x_1^2 + x_2^2 \leq 5 \tag{5.3.5.2}$$

$$x_1 \geq 0, x_2 \geq 0 \tag{5.3.5.3}$$

5.3.6 Solve

$$\min_{\mathbf{x}} \quad x_1 + x_2 \tag{5.3.6.1}$$

with the constraints

$$x_1^2 - x_1 + x_2^2 \leq 0 \tag{5.3.6.2}$$

where $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$

Solution:

Graphical solution:

```
codes/kkt/2.15.py
```

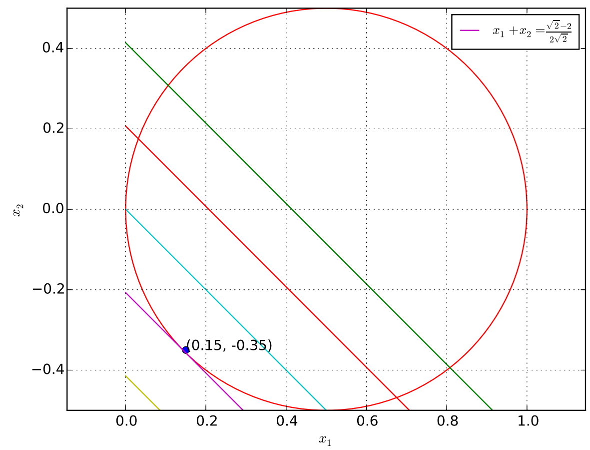

Fig. 5.3.6. Optimal solution is the lower tangent to the circle

6 Quadratic Programming

6.1. An apache helicopter of the enemy is flying along the curve given by

$$y = x^2 + 7 \tag{6.1.1}$$

A soldier, placed at

$$\mathbf{P} = \begin{pmatrix} 3 \\ 7 \end{pmatrix}. \tag{6.1.2}$$

wants to shoot the heicopter when it is nearest to him. Express this as an optimization problem.

Solution: The given problem can be expressed as

$$\min_{\mathbf{x}} \|\mathbf{x} - \mathbf{P}\|^2 \tag{6.1.3}$$

$$\text{s.t. } \mathbf{x}^T \mathbf{V} \mathbf{x} + \mathbf{u}^T \mathbf{x} + d = 0 \tag{6.1.4}$$

where

$$\mathbf{V} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \tag{6.1.5}$$

$$\mathbf{u} = -\begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{6.1.6}$$

$$d = 7 \tag{6.1.7}$$

6.2. Show that the constraint in 6.1.3 is nonconvex.

6.3. Show that the following *relaxation* makes (6.1.3) a convex optimization problem.

$$\min_{\mathbf{x}} (\mathbf{x} - \mathbf{P})^T (\mathbf{x} - \mathbf{P}) \tag{6.3.1}$$

$$\text{s.t. } \mathbf{x}^T \mathbf{V} \mathbf{x} + \mathbf{u}^T \mathbf{x} \leq 0 \tag{6.3.2}$$

6.4. Solve (6.3.1) using cvxpy.
Solution: The following code yields the minimum distance as 2.236 and the nearest point on the curve as

$$\mathbf{Q} = \begin{pmatrix} 1 \\ 8 \end{pmatrix} \tag{6.4.1}$$

```
codes/opt/qp_cvx.py
```

6.5. Solve (6.3.1) using the method of Lagrange multipliers.

6.6. Graphically verify the solution to Problem 6.1.
Solution: The following code plots Fig. 6.6

```
codes/opt/qp_parab.py
```

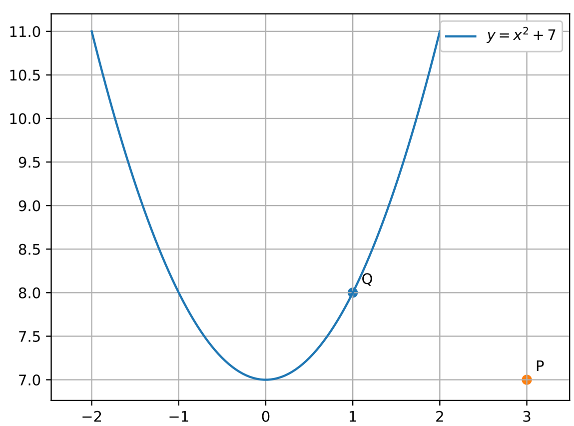

Fig. 6.6. **Q** is closest to **P**

.

6.7. Solve (6.3.1) using gradient descent.

7 Semi-definite Programming

7.1 Solve

$$\min_{\mathbf{x}} \quad x_1 + x_2 \tag{7.1.1}$$

with the constraints

$$x_1^2 - x_1 + x_2^2 \le 0 \tag{7.1.2}$$

where $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$

7.2 The problem

$$\min_{\mathbf{X}} x_{11} + x_{12} \tag{7.2.1}$$

with constraints

$$x_{11} + x_{22} = 1 \tag{7.2.2}$$

$$\mathbf{X} \succeq 0 \quad (\succeq \text{ means positive definite}) \tag{7.2.3}$$

where

$$\mathbf{X} = \begin{pmatrix} x_{11} & x_{12} \\ x_{12} & x_{22} \end{pmatrix} \tag{7.2.4}$$

is known as a semi-definite program.

7.3 Frame Problem 7.2 in terms of matrices.
Solution: It is easy to verify that

$$x_{11} + x_{12} = \begin{pmatrix} 1 & 1 \end{pmatrix} \mathbf{X}^T \begin{pmatrix} 1 \\ 0 \end{pmatrix} \tag{7.3.1}$$

and

$$x_{11} + x_{22} = \begin{pmatrix} 1 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \mathbf{X} & \mathbf{0} \\ \mathbf{0} & \mathbf{X} \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix} \tag{7.3.2}$$

Thus, Problem 7.2 can be expressed as

$$\min_{\mathbf{X}} \begin{pmatrix} 1 & 1 \end{pmatrix} \mathbf{X}^T \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad s.t$$
$$\begin{pmatrix} 1 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \mathbf{X} & \mathbf{0} \\ \mathbf{0} & \mathbf{X} \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix} = 1, \tag{7.3.3}$$
$$\mathbf{X} \succeq 0$$

7.4 Solve (7.3.3) using *cvxpy*. Compare with the solution in problem 7.1.
Solution:

```
wget https://raw.
    githubusercontent.com/
    gadepall/optimization/master
    /manual/codes/3.1-cvx.py
```

7.5 Minimize

$$-x_{11} - 2x_{12} - 5x_{22} \tag{7.5.1}$$

subject to

$$2x_{11} + 3x_{12} + x_{22} = 7 \tag{7.5.2}$$
$$x_{11} + x_{12} \ge 1 \tag{7.5.3}$$
$$x_{11}, x_{12}, x_{22} \ge 0 \tag{7.5.4}$$
$$\begin{pmatrix} x_{11} & x_{12} \\ x_{12} & x_{22} \end{pmatrix} \succeq 0 \tag{7.5.5}$$

using *cvxpy*.

7.6 Repeat the above exercise by converting the problem into a convex optimization problem in two variables and using graphical plots.
7.7 Solve the above problem using the KKT conditions. Comment.

8 Linear Programming

8.1. Solve

$$\max_{\mathbf{x}} Z = \begin{pmatrix} 4 & 1 \end{pmatrix} \mathbf{x} \tag{8.1.1}$$

$$s.t. \quad \begin{pmatrix} 1 & 1 \\ 3 & 1 \end{pmatrix} \mathbf{x} \le \begin{pmatrix} 50 \\ 90 \end{pmatrix} \tag{8.1.2}$$

$$\mathbf{x} \ge \mathbf{0} \tag{8.1.3}$$

using cvxpy.
Solution: The given problem can be expressed in general as

$$\max_{\mathbf{x}} \mathbf{c}^T \mathbf{x} \tag{8.1.4}$$

$$s.t. \quad \mathbf{A}\mathbf{x} \le \mathbf{b}, \tag{8.1.5}$$

$$\mathbf{x} \ge \mathbf{0} \tag{8.1.6}$$

where

$$\mathbf{c} = \begin{pmatrix} 4 \\ 1 \end{pmatrix} \tag{8.1.7}$$

$$\mathbf{A} = \begin{pmatrix} 1 & 1 \\ 3 & 1 \end{pmatrix} \tag{8.1.8}$$

$$\mathbf{b} = \begin{pmatrix} 50 \\ 90 \end{pmatrix} \tag{8.1.9}$$

and can be solved using *cvxpy* through the following code

```
codes/opt/lp_cvx.py
```

to obtain

$$\mathbf{x} = \begin{pmatrix} 30 \\ 0 \end{pmatrix}, Z = 120 \tag{8.1.10}$$

8.2. Graphically, show that the feasible region in Problem 8.1 result in the interior of a convex polygon and the optimal point is one of the vertices. **Solution:** The following code plots Fig. 8.2.

```
codes/opt/lp_cvx.py
```

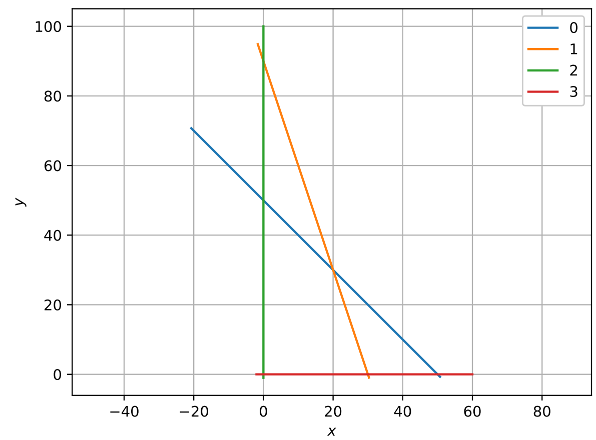

Fig. 8.2.

8.3. Solve

$$\min_{\mathbf{x}} Z = \begin{pmatrix} 3 & 9 \end{pmatrix} \mathbf{x} \tag{8.3.1}$$

$$s.t. \quad \begin{pmatrix} 1 & 3 \\ -1 & -1 \\ 1 & -1 \end{pmatrix} \mathbf{x} \le \begin{pmatrix} 60 \\ -10 \\ 0 \end{pmatrix} \tag{8.3.2}$$

$$\mathbf{x} \ge \mathbf{0} \tag{8.3.3}$$

Solution: The following code

```
codes/opt/lp_cvx_mult.py
```

is used to obtain

$$\mathbf{x} = \begin{pmatrix} 15 \\ 15 \end{pmatrix}, Z = 180 \tag{8.3.4}$$

8.4. Solve

$$\min_{\mathbf{x}} Z = \begin{pmatrix} -50 & 20 \end{pmatrix} \mathbf{x} \tag{8.4.1}$$

$$s.t. \quad \begin{pmatrix} -2 & 1 \\ -3 & -1 \\ 2 & -3 \end{pmatrix} \mathbf{x} \le \begin{pmatrix} 5 \\ -3 \\ 12 \end{pmatrix} \tag{8.4.2}$$

$$\mathbf{x} \ge \mathbf{0} \tag{8.4.3}$$

Solution: The following code

```
codes/opt/lp_cvx_nosol.py
```

shows that the given problem has no solution.

8.5. Verify all the above solutions using Lagrange multipliers.
8.6. Repeat the above exercise using the Simplex method.
8.7. **(Diet problem)**: A dietician wishes to mix two types of foods in such a way that vitamin contents of the mixture contain atleast 8 units of

vitamin A and 10 units of vitamin C. Food 'I' contains 2 units/kg of vitamin A and 1 unit/kg of vitamin C. Food 'II' contains 1 unit/kg of vitamin A and 2 units/kg of vitamin C. It costs Rs 50 per kg to purchase Food 'I' and Rs 70 per kg to purchase Food 'II'. Formulate this problem as a linear programming problem to minimise the cost of such a mixture.

Solution: Let the mixture contain x kg of food I and y kg of food II.

The given problem can be expressed as

Resources	Food		Requirement
	I	II	
Vitamin A	2	1	Atleast 8 Units
Vitamin C	1	2	Atleast 10 Units
Cost	50	70	

$$\min_{\mathbf{x}} Z = \begin{pmatrix} 50 & 70 \end{pmatrix} \mathbf{x} \tag{8.7.1}$$

$$s.t. \quad \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} \mathbf{x} \geq \begin{pmatrix} 8 \\ 10 \end{pmatrix} \tag{8.7.2}$$

$$\mathbf{x} \geq \mathbf{0} \tag{8.7.3}$$

The corner points of the feasible region are available in Table 8.7 and plotted in Fig. 8.7.

Corner Point	$Z = 50x + 70y$
(0,8)	560
(2,4)	380
(10,0)	500

TABLE 8.7

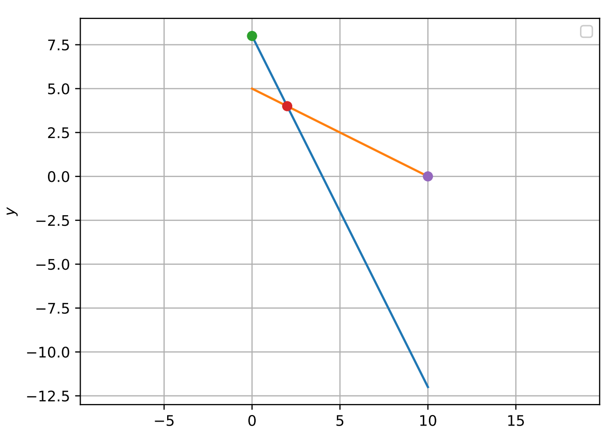

Fig. 8.7.

The smallest value of Z is 380 at the point (2,4). But the feasible region is unbounded therefore we draw the graph of the inequality

$$50x + 70y < 380 \tag{8.7.4}$$

to check whether the resulting open half has any point common with the feasible region but on checking it doesn't have any points in common. Thus the minimum value of Z is 380 attained at $\begin{pmatrix} 2 \\ 4 \end{pmatrix}$. Hence optimal mixing strategy for the dietician would be to mix 2 Kg of Food I and 4 Kg of Food II. The following code provides the solution to (8.7.3).

```
codes/opt/diet.py
```

8.8. **(Allocation problem)** A cooperative society of farmers has 50 hectare of land to grow two crops X and Y. The profit from crops X and Y per hectare are estimated as Rs 10,500 and Rs 9,000 respectively. To control weeds, a liquid herbicide has to be used for crops X and Y at rates of 20 litres and 10 litres per hectare. Further, no more than 800 litres of herbicide should be used in order to protect fish and wild life using a pond which collects drainage from this land. How much land should be allocated to each crop so as to maximise the total profit of the society?

Solution: The given problem can be formulated as

$$\max_{\mathbf{x}} Z = \begin{pmatrix} 10500 & 9000 \end{pmatrix} \mathbf{x} \tag{8.8.1}$$

$$s.t. \quad \begin{pmatrix} 20 & 10 \end{pmatrix} \mathbf{x} \leq 800 \tag{8.8.2}$$

$$\begin{pmatrix} 1 & 1 \end{pmatrix} \mathbf{x} = 50 \tag{8.8.3}$$

Fig 8.8 shows the intersection of various lines and the optimal point as indicated.

The following code provides the solution to (8.8.3) at $\begin{pmatrix} 30 \\ 20 \end{pmatrix}$.

```
codes/opt/allocation.py
```

8.9. **(Manufacturing problem)** A manufacturer has three machines I, II and III installed in his factory. Machines I and II are capable of being operated for at most 12 hours whereas machine III must be operated for atleast 5 hours a day. She produces only two items M and N each requiring the use of all the three machines. The

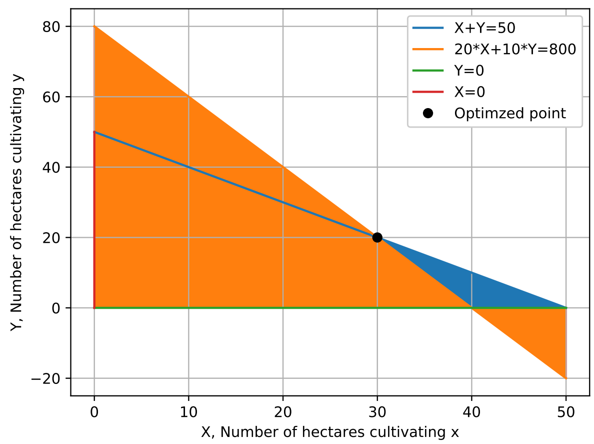

Fig. 8.8. Feasible region for allocation Problem

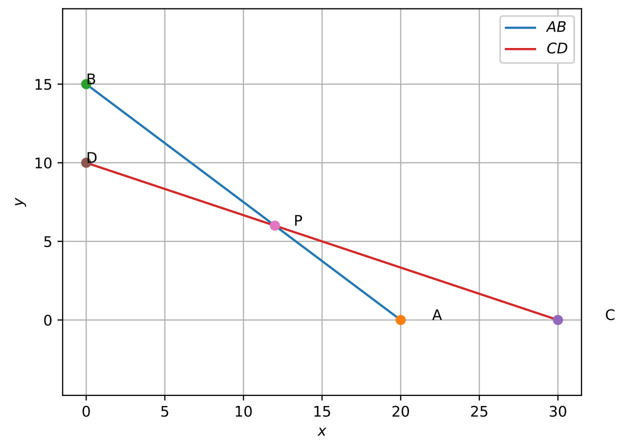

Fig. 8.9. Feasible region for manufacturing Problem

Fig. 8.8.

number of hours required for producing 1 unit of each of M and N on the three machines are given in Table 8.9

Number of hours required on machines			
Items	**I**	**II**	**III**
M	1	2	1
N	2	1	1.25

TABLE 8.9

She makes a profit of Rs 600 and Rs 400 on items M and N respectively. How many of each item should she produce so as to maximise her profit assuming that she can sell all the items that she produced? What will be the maximum profit?

Solution: The given problem can be formulated as

$$\max_{\mathbf{x}} Z = \begin{pmatrix} 80000 & 12000 \end{pmatrix} \mathbf{x} \tag{8.9.1}$$

$$s.t. \quad \begin{pmatrix} 3 & 4 \\ 1 & 3 \end{pmatrix} \mathbf{x} \leq \begin{pmatrix} 60 \\ 30 \end{pmatrix} \tag{8.9.2}$$

Fig 8.9 shows the intersection of various lines and the optimal point as indicated.

The following code provides the solution to (8.9.2) at $\begin{pmatrix} 12 \\ 6 \end{pmatrix}$.

```
codes/opt/Manufacturing.py
```

8.10. **(Transportation problem)** There are two factories located one at place P and the other at place Q. From these locations, a certain commodity is to be delivered to each of the three depots situated at A, B and C. The weekly requirements of the depots are respectively 5, 5 and 4 units of the commodity while the production capacity of the factories at P and Q are respectively 8 and 6 units. The cost of transportation per unit is given below where A,B,C are cost in ruppes:

From/To	A	B	C
P	160	100	150
Q	100	120	100

How many units should be transported from each factory to each depot in order that the transportation cost is minimum. What will be the minimum transportation cost?

Solution: The given problem can be formulated as

$$\min_{\mathbf{x}} Z = \begin{pmatrix} 10 & -70 \end{pmatrix} \mathbf{x} \tag{8.10.1}$$

$$s.t. \quad \begin{pmatrix} 1 & 1 \\ -1 & -1 \end{pmatrix} \mathbf{x} \leq \begin{pmatrix} 8 \\ -4 \end{pmatrix} \tag{8.10.2}$$

$$\mathbf{x} \leq \begin{pmatrix} 5 \\ 5 \end{pmatrix} \tag{8.10.3}$$

Fig. 8.10 shows the intersection of various lines and the optimal point indicated as OPT PT.

The following code provides the solution to (8.10.3) at $\begin{pmatrix} 0 \\ 5 \end{pmatrix}$.

```
codes/opt/Transportation.py
```

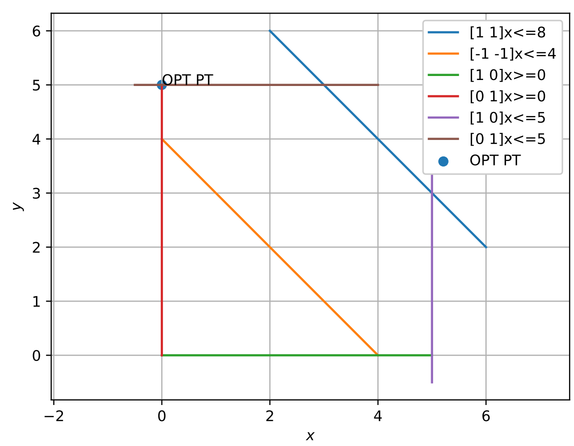

Fig. 8.10. Feasible region for Transportation Problem

9 Geometric Programming

9.1. Amongst all open (from the top) right circular cylindrical boxes of volume $125\pi\ cm^3$, find the dimensions of the box which has the least surface area.

Solution: Let r be the radius of the cylinder and h be the height. Then the surface area is

$$S = \pi r^2 + 2\pi rh \tag{9.1.1}$$

Also, the volume is

$$V = \pi r^2 h \tag{9.1.2}$$

9.2. The given problem can then be formulated as

$$S = \min_{r,h} \pi r^2 + 2\pi rh \tag{9.2.1}$$

$$\text{s.t} \quad \pi r^2 h = 125 \tag{9.2.2}$$

which is a *disciplined geometric programming* (DGP) problem that can be solved using *cvxpy*. DGP is a subset of *log-log-convex program* (LLCP). An LLCP is defined as

$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & f_i(x) \le \tilde{f}_i, \quad i = 1, \dots, m \\ & g_i(x) = \tilde{g}_i, \quad i = 1, \dots, p, \end{array} \tag{9.2.3}$$

where the functions f_i are log-log convex, $\tilde{f}_i$ are log-log concave, and the functions g_i and $\tilde{g}_i$ are log-log affine. An optimization problem with constraints of the above form in which the goal is to maximize a log-log concave function is also an LLCP. A function

$$f : D \subseteq \mathbf{R}^n_{++} \to \mathbf{R} \tag{9.2.4}$$

is said to be log-log convex if the function

$$F(u) = \log f(e^u) \tag{9.2.5}$$

with domain

$$\{u \in \mathbf{R}^n : e^u \in D\} \tag{9.2.6}$$

is convex (where $\mathbf{R}^n_{++}$ denotes the set of positive reals and the logarithm and exponential are meant elementwise); the function F is called the log-log transformation of f. The function f is log-log concave if F is concave, and it is log-log affine if F is affine. LLCPs are problems that become convex after the variables, objective functions, and constraint functions are replaced with their logs, an operation that we refer to as a log-log transformation. LLCPs generalize geometric programming.

9.3. Alternatively, from (9.1.1) and (9.1.2)

$$S(r) = \pi r^2 + \frac{2V}{r} \tag{9.3.1}$$

$$\implies S'(r) = 2\pi r - \frac{2V}{r^2} \tag{9.3.2}$$

$$\text{and } S''(r) = 2\pi + \frac{4V}{r^3} > 0 \tag{9.3.3}$$

Thus, $S(r)$ has a minimum which can be obtained from (9.3.2) as

$$2\pi r - \frac{2V}{r^2} = 0 \tag{9.3.4}$$

$$\implies r = \left(\frac{V}{\pi}\right)^{\frac{1}{3}} \tag{9.3.5}$$

$$= 5 \quad \text{and} \tag{9.3.6}$$

$$h = \frac{V}{\pi r^2} = 5 \tag{9.3.7}$$

upon substituting numerical values. This is verified in Fig. 9.3.

9.4. Using gradient descent, the update equation can be expressed as

$$r_{n+1} = r_n - \gamma S'(r_n) \tag{9.4.1}$$

where $r_0 = 2$ and $\gamma = 0.001$ are chosen by the user. These values need to be suitably guessed for the algorithm to converge.

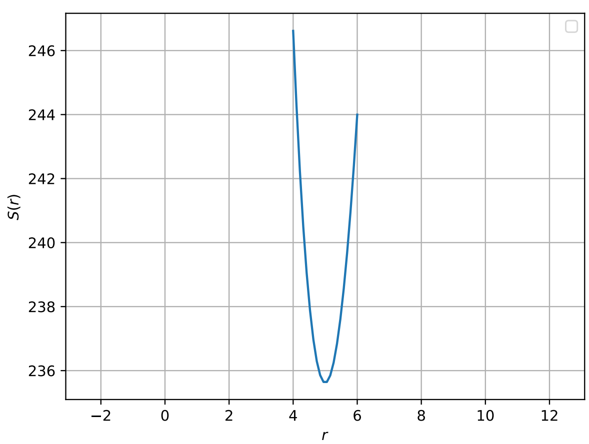

Fig. 9.3.