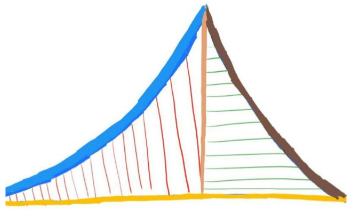
OPTIMIZATION

Through High School Math

G. V. V. Sharma



Contents

Introduction

This book introduces optimization through high school math.

Chapter 1

Distance from a Line

1.1 Reduce $x - \sqrt{3}y + 8 = 0$ into normal form. Find its perpendicular distance from the origin and angle between perpendicular and the positive x-axis.

Solution:

(a) Optimization Problem

Let $\mathbf{O}$ be the point from where we have to find the perpendicular distance and $\mathbf{P}$ be the foot of the perpendicular. The optimization problem can be expressed as

$$\min_{\mathbf{x}} \|\mathbf{x} - \mathbf{O}\|^2 \tag{1.1.1}$$

$$\text{s.t.} \quad \mathbf{n}^T\mathbf{x} = c \tag{1.1.2}$$

where

$$\mathbf{n} = \begin{pmatrix} -1 \\ \sqrt{3} \end{pmatrix}, c = 8 \tag{1.1.3}$$

The line equation can be expressed as

$$\mathbf{x} = \mathbf{A} + \lambda\mathbf{m} \tag{1.1.4}$$

where

$$\mathbf{m} = \begin{pmatrix} 1 \\ \frac{1}{\sqrt{3}} \end{pmatrix}, \mathbf{A} = \begin{pmatrix} -8 \\ 0 \end{pmatrix} \tag{1.1.5}$$

i. Using the parameric form, Substituting (1.1.4) in (1.1.1), the optimization problem becomes

$$\min_{\lambda} \left\|\lambda\mathbf{m} + (\mathbf{A} - \mathbf{O})\right\|^2 \tag{1.1.6}$$

$$\implies \min_{\lambda} f(\lambda) = \lambda^2 \left\|\mathbf{m}\right\|^2 + 2\lambda \left(\mathbf{A} - \mathbf{O}\right)^\top \mathbf{m} + + \left\|\mathbf{A} - \mathbf{O}\right\|^2 \tag{1.1.7}$$

$\because$ the coefficient of $\lambda^2 > 0$, (1.1.7) is a convex function. Thus,

$$f''(\lambda) = 2\left\|\mathbf{m}\right\|^2 \tag{1.1.8}$$

$$\because f''(\lambda) > 0, f'(\lambda_{min}) = 0, \text{ for } \lambda_{min} \tag{1.1.9}$$

yielding

$$f'(\lambda_{min}) = 2\lambda_{min} \left\|\mathbf{m}\right\|^2 + 2\left(\mathbf{A} - \mathbf{O}\right)^\top \mathbf{m} = 0 \tag{1.1.10}$$

$$\lambda_{min} = -\frac{\left(\mathbf{A} - \mathbf{O}\right)^\top \mathbf{m}}{\left\|\mathbf{m}\right\|^2} \tag{1.1.11}$$

We choose

$$\mathbf{O} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \tag{1.1.12}$$

Substituting the values of $\mathbf{A}$, $\mathbf{O}$ and $\mathbf{m}$ in equation (1.1.11)

$$\lambda_{min} = -\frac{\left(\begin{pmatrix} -8 \\ 0 \end{pmatrix} - \begin{pmatrix} 0 \\ 0 \end{pmatrix}\right)^\top \begin{pmatrix} 1 \\ \frac{1}{\sqrt{3}} \end{pmatrix}}{\left\| \begin{pmatrix} 1 \\ \frac{1}{\sqrt{3}} \end{pmatrix} \right\|^2} \tag{1.1.13}$$

$$= 6 \tag{1.1.14}$$

Substituring this value in equation (1.1.4)

$$\mathbf{x}_{min} = \mathbf{P} = \begin{pmatrix} -8 \\ 0 \end{pmatrix} + 6 \begin{pmatrix} 1 \\ \frac{1}{\sqrt{3}} \end{pmatrix} \tag{1.1.15}$$

$$= \begin{pmatrix} -2 \\ 2\sqrt{3} \end{pmatrix} \tag{1.1.16}$$

$$OP = \|\mathbf{P} - \mathbf{O}\|^2 \tag{1.1.17}$$

$$= 4 \tag{1.1.18}$$

ii. Solving using cvxpy, with

$$\mathbf{n} = \begin{pmatrix} 1 \\ -\sqrt{3} \end{pmatrix} \tag{1.1.19}$$

$$\mathbf{O} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \tag{1.1.20}$$

$$c = -8 \tag{1.1.21}$$

$$\min_{\mathbf{x}} \|\mathbf{x} - \mathbf{O}\|^2 = 4, \mathbf{x}_{min} = \begin{pmatrix} -2 \\ 3.46 \end{pmatrix} \tag{1.1.22}$$

The relevant figures are shown in 1.1.1 and 1.1.2

(b) Gradient Descent

The given line can be represented in parametric form as

$$\mathbf{x} = \mathbf{A} + \lambda \mathbf{m} \tag{1.1.23}$$

where

$$\mathbf{A} = \begin{pmatrix} -8 \\ 0 \end{pmatrix} \tag{1.1.24}$$

$$\mathbf{O} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \tag{1.1.25}$$

$$\mathbf{m} = \begin{pmatrix} 1 \\ \frac{1}{\sqrt{3}} \end{pmatrix} \tag{1.1.26}$$

$$\tag{1.1.27}$$

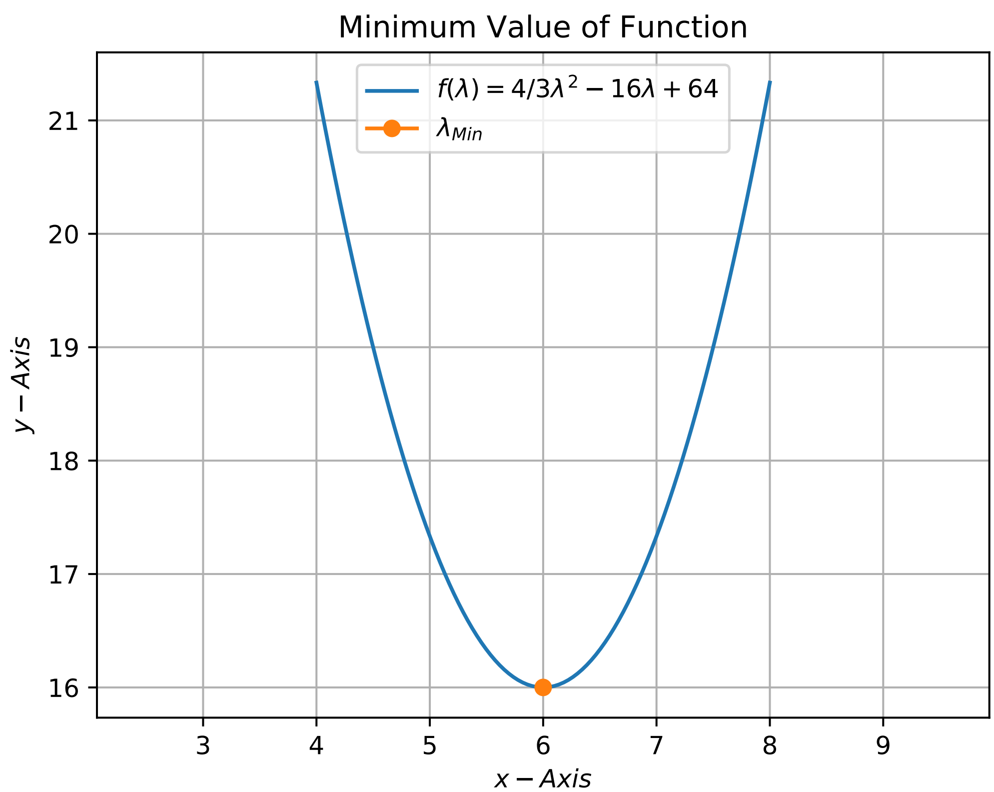


Figure 1.1.1:

yielding

$$f(\lambda) = \frac{4}{3}\lambda^2 - 16\lambda + 64 \tag{1.1.28}$$

$$f'(\lambda) = \frac{8}{3}\lambda - 16 \tag{1.1.29}$$

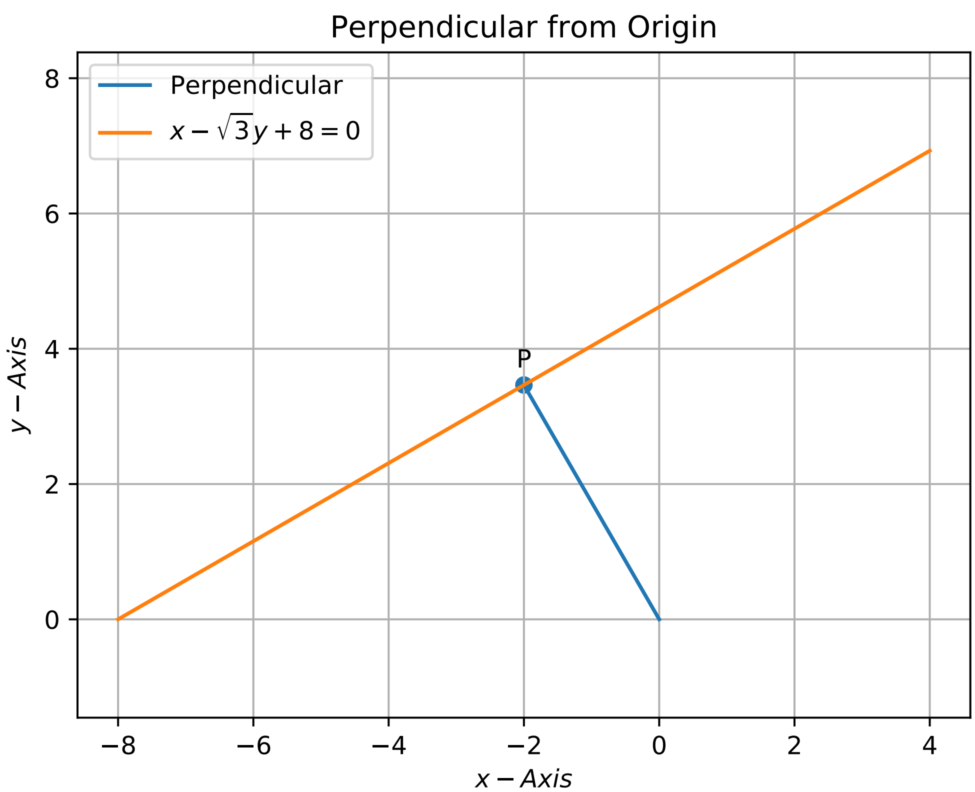


Figure 1.1.2:

Computing λ_{min} using Gradient Descent method:

$$\lambda_{n+1} = \lambda_n - \alpha f'(\lambda_n) \tag{1.1.30}$$

$$\lambda_{n+1} = \lambda_n \left(1 - \frac{8}{3}\alpha\right) + 16\alpha \tag{1.1.31}$$

Taking the one-sided Z-transform on both sides of (1.1.31),

$$z\Lambda(z) = \left(1 - \frac{8}{3}\alpha\right)\Lambda(z) + \frac{16\alpha}{1 - z^{-1}} \tag{1.1.32}$$

$$\Lambda(z) = \frac{16\alpha z^{-1}}{\left(1 - z^{-1}\right)\left(1 - \left(1 - \frac{8}{8}\alpha\right)z^{-1}\right)} \tag{1.1.33}$$

$$= 6\left(\frac{1}{1 - z^{-1}} - \frac{1}{1 - \left(1 - \frac{8}{3}\alpha\right)z^{-1}}\right) \tag{1.1.34}$$

$$= 6\sum_{k=0}^{\infty}\left(1 - \left(1 - \frac{8}{3}\alpha\right)^{k}\right)z^{-k} \tag{1.1.35}$$

From (1.1.35), the ROC is

$$|z| > \max\left\{1, \left|1 - \frac{8}{3}\alpha\right|\right\} \tag{1.1.36}$$

$$\implies -1 < \left|1 - \frac{8}{3}\alpha\right| < 1 \tag{1.1.37}$$

$$\implies 0 < \alpha < \frac{3}{4} \tag{1.1.38}$$

Thus, if α satisfies (1.1.38), then from (1.1.35),

$$\lim_{n\to\infty}\lambda_n = 6 \tag{1.1.39}$$

Choosing

i. $\alpha = 0.001$

ii. precision = 0.0000001

iii. n = 10000000

iv. λ_0 = -5

$$\lambda_{min} = 6 \tag{1.1.40}$$

Substituting the values of $\mathbf{A}$, $\mathbf{m}$ and λ_{min} in equation (1.1.23)

$$\mathbf{x}_{min} = \mathbf{P} = \begin{pmatrix} -8 \\ 0 \end{pmatrix} + 6 \begin{pmatrix} 1 \\ \frac{1}{\sqrt{3}} \end{pmatrix} \tag{1.1.41}$$

$$= \begin{pmatrix} -8 \\ 0 \end{pmatrix} + \begin{pmatrix} 6 \\ \frac{6}{\sqrt{3}} \end{pmatrix} \tag{1.1.42}$$

$$= \begin{pmatrix} -2 \\ 2\sqrt{3} \end{pmatrix} \tag{1.1.43}$$

$$OP = \left\|\mathbf{P} - \mathbf{O}\right\|^2 \tag{1.1.44}$$

$$= \left\| \begin{pmatrix} -2 \\ 2\sqrt{3} \end{pmatrix} - \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right\| \tag{1.1.45}$$

$$= \sqrt{2^2 + 12} = \sqrt{16} = 4 \tag{1.1.46}$$

See Figs. 1.1.3 and 1.1.4.

(c) Lagrange Multipliers

The given problem can be formulated as

$$\min_{\mathbf{x}} f\left(\mathbf{x}\right) = \left\|\mathbf{x} - \mathbf{O}\right\|^2 \tag{1.1.47}$$

$$\text{s.t.} \quad g\left(\mathbf{x}\right) = \mathbf{n}^T\mathbf{x} - c = 0 \tag{1.1.48}$$

where

$$\mathbf{n} = \begin{pmatrix} 1 \\ -\sqrt{3} \end{pmatrix}, \mathbf{O} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \text{ and } c = -8 \tag{1.1.49}$$

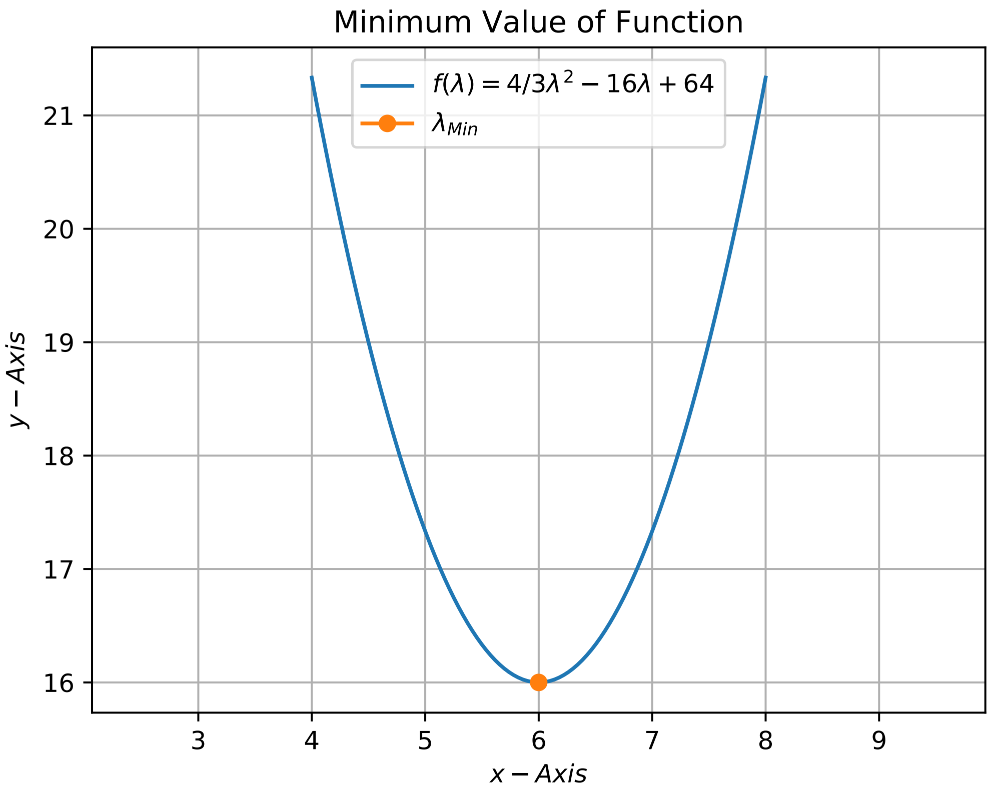


Figure 1.1.3:

Define

$$H(\mathbf{x}, \lambda) = f(\mathbf{x}) - \lambda g(\mathbf{x}) \tag{1.1.50}$$

and we find that

$$\nabla f(\mathbf{x}) = 2(\mathbf{x} - \mathbf{O}) \tag{1.1.51}$$

$$\nabla g(\mathbf{x}) = \mathbf{n} \tag{1.1.52}$$

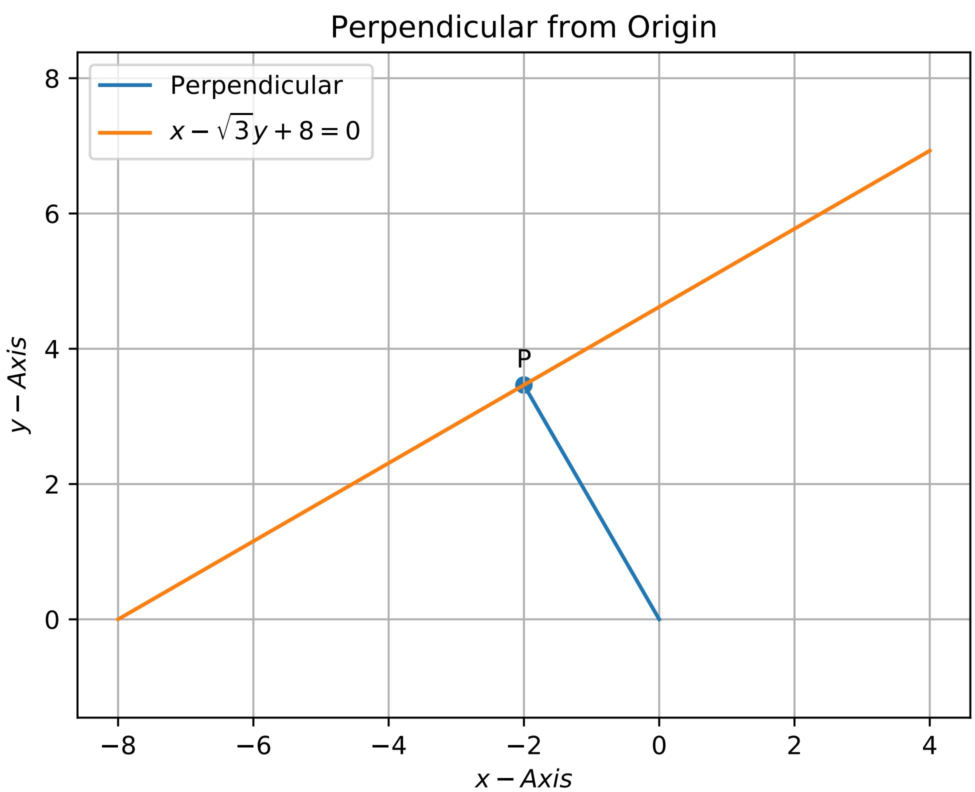


Figure 1.1.4:

We have to find $\lambda \in \mathbb{R}$ such that

$$\nabla H(\mathbf{x}, \lambda) = 0 \tag{1.1.53}$$

$$\implies 2(\mathbf{x} - \mathbf{O}) - \lambda \mathbf{n} = 0 \tag{1.1.54}$$

$$\implies \mathbf{x} = \frac{\lambda}{2}\mathbf{n} + \mathbf{O} \tag{1.1.55}$$

Substituting (1.1.55) in (1.1.48)

$$\mathbf{n}^\top \left(\frac{\lambda}{2}\mathbf{n} + \mathbf{O}\right) - c = 0 \tag{1.1.56}$$

$$\implies \lambda = \frac{2\left(c - \mathbf{n}^\top \mathbf{O}\right)}{\|\mathbf{n}\|^2} \tag{1.1.57}$$

Substituting the value of λ in (1.1.54),

$$\mathbf{x}_{min} = \mathbf{P} = \mathbf{O} + \frac{\mathbf{n}\left(c - \mathbf{n}^\top \mathbf{O}\right)}{\|\mathbf{n}\|^2} \tag{1.1.58}$$

$$= \begin{pmatrix} 0 \\ 0 \end{pmatrix} + \frac{\begin{pmatrix} 1 \\ -\sqrt{3} \end{pmatrix}\left(-8 - \begin{pmatrix} 1 & -\sqrt{3} \end{pmatrix}\begin{pmatrix} 0 \\ 0 \end{pmatrix}\right)}{4} \tag{1.1.59}$$

$$= \begin{pmatrix} -2 \\ 2\sqrt{3} \end{pmatrix} \tag{1.1.60}$$

$$OP = \|\mathbf{P} - \mathbf{O}\|^2 \tag{1.1.61}$$

$$= \left\| \begin{pmatrix} -2 \\ 2\sqrt{3} \end{pmatrix} - \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right\| \tag{1.1.62}$$

$$= 4 \tag{1.1.63}$$

The relevant figure is shown in 1.1.5

1.2 Reduce the equation $y - 2 = 0$ into normal form. Find the perpendicular distances from the origin and angle between perpendicular and the positive x-axis.

Solution:

(a) Optimization Problem

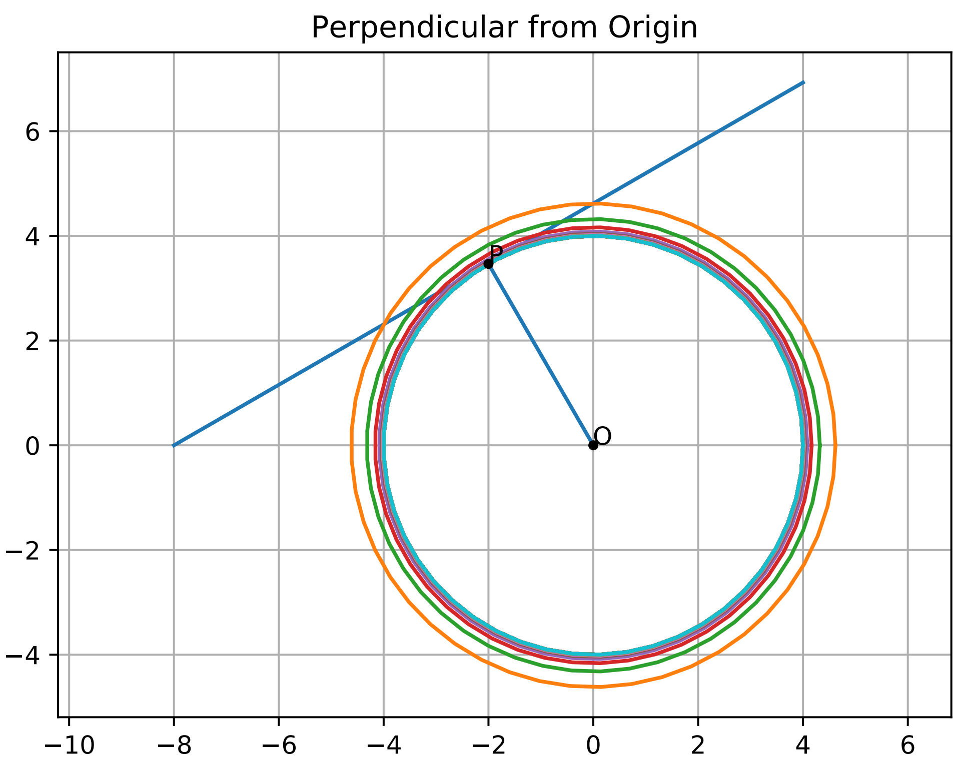


Figure 1.1.5:

The given equation can be written as

$$\begin{pmatrix} 0 & 1 \end{pmatrix} \mathbf{x} = 2 \tag{1.2.1}$$

$$\implies \mathbf{n} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \mathbf{m} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \tag{1.2.2}$$

Equation (1.2.1) can be represented in parametric form as

$$\mathbf{x} = \mathbf{A} + \lambda \mathbf{m} \tag{1.2.3}$$

where

$$\mathbf{A} = \begin{pmatrix} 2 \\ 2 \end{pmatrix}. \tag{1.2.4}$$

Let $\mathbf{O}$ be the origin. The perpendicular distance will be the minimum distance from $\mathbf{O}$ to the line. Let $\mathbf{P}$ be the foot of perpendicular. This problem can be formulated as an optimization problem as

$$d = \min_{\mathbf{x}} \left\|\mathbf{x} - \mathbf{O}\right\|^2 \tag{1.2.5}$$

$$= \min_{\lambda} \left\|\mathbf{A} + \lambda \mathbf{m} - \mathbf{O}\right\|^2 \tag{1.2.6}$$

$$= f(\lambda) = \left\|\mathbf{m}\right\|^2 \lambda^2 + 2\mathbf{A}^\top \mathbf{m} + \left\|\mathbf{A}\right\|^2 \tag{1.2.7}$$

$$= \lambda^2 + 4\lambda + 8 \tag{1.2.8}$$

$\because$ the coefficient of $\lambda^2 > 0$, (1.2.7) is convex.

$$f'(\lambda) = 2\left\|\mathbf{m}\right\|^2 \lambda + \left(\mathbf{A}^\top \mathbf{m} + \mathbf{m}^\top \mathbf{A}\right) \tag{1.2.9}$$

i. Computing λ_{min} using Derivative method

$$f''(\lambda) = 2 \tag{1.2.10}$$

$$\because f''(\lambda) > 0, f'(\lambda_{min}) = 0, \text{ for } \lambda_{min} \tag{1.2.11}$$

$$f'(\lambda_{min}) = 2\|\mathbf{m}\|^2 \lambda_{min} + \left(\mathbf{A}^\top \mathbf{m} + \mathbf{m}^\top \mathbf{A}\right) \tag{1.2.12}$$

$$\therefore \lambda_{min} = -\frac{\left(\mathbf{A}^\top \mathbf{m} + \mathbf{m}^\top \mathbf{A}\right)}{2\|\mathbf{m}\|^2} = -2 \tag{1.2.13}$$

Thus,

$$\mathbf{x}_{min} = \mathbf{P} = \begin{pmatrix} 2 \\ 2 \end{pmatrix} + (-2)\begin{pmatrix} 1 \\ 0 \end{pmatrix} \tag{1.2.14}$$

$$= \begin{pmatrix} 0 \\ 2 \end{pmatrix} \tag{1.2.15}$$

$$OP = \|\mathbf{P} - \mathbf{O}\| \tag{1.2.16}$$

$$= \left\| \begin{pmatrix} 0 \\ 2 \end{pmatrix} - \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right\| \tag{1.2.17}$$

$$= 2 \tag{1.2.18}$$

ii. Solving using cvxpy, with

$$\mathbf{n} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{1.2.19}$$

$$\mathbf{O} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \tag{1.2.20}$$

$$c = 2 \tag{1.2.21}$$

$$\min_{\mathbf{x}} \|\mathbf{x} - \mathbf{O}\|^2 = 2, \mathbf{x}_{min} = \begin{pmatrix} 0 \\ 2 \end{pmatrix} \tag{1.2.22}$$

See Figs. 1.2.1 and 1.2.2.

(b) Gradient Descent

The given equation can be represented in parametric form as

$$\mathbf{x} = \mathbf{A} + \lambda \mathbf{m} \tag{1.2.23}$$

where

$$\mathbf{A} = \begin{pmatrix} 2 \\ 2 \end{pmatrix}, \mathbf{m} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \tag{1.2.24}$$

Let **O** be the origin. The perpendicular distance will be the minimum distance from **O** to the line. Let **P** be the foot of perpendicular. This problem can be

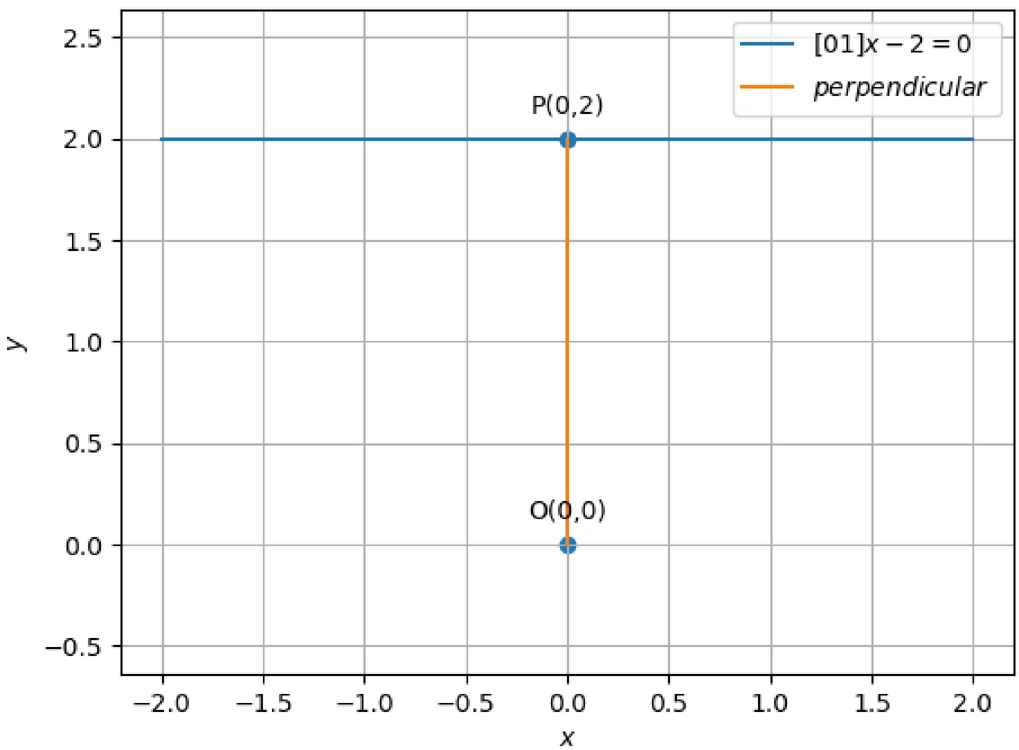


Figure 1.2.1:

formulated as an optimization problem as follows:

$$d_{\min} = \min_{\mathbf{x}} \left\|\mathbf{x} - \mathbf{O}\right\|^2 \tag{1.2.25}$$

$$= \min_{\lambda} \left\|\mathbf{A} + \lambda\mathbf{m} - \mathbf{O}\right\|^2 \tag{1.2.26}$$

$$= \left\|\mathbf{m}\right\|^2 \lambda^2 + 2\lambda\mathbf{A}^\top\mathbf{m} + \left\|\mathbf{A}\right\|^2 \tag{1.2.27}$$

$$= \lambda^2 + 4\lambda + 8 = f'\left(\lambda\right) \tag{1.2.28}$$

upon substituting numerical values. $\because$ the coefficient of $\lambda^2 > 0$, (1.2.27) is a

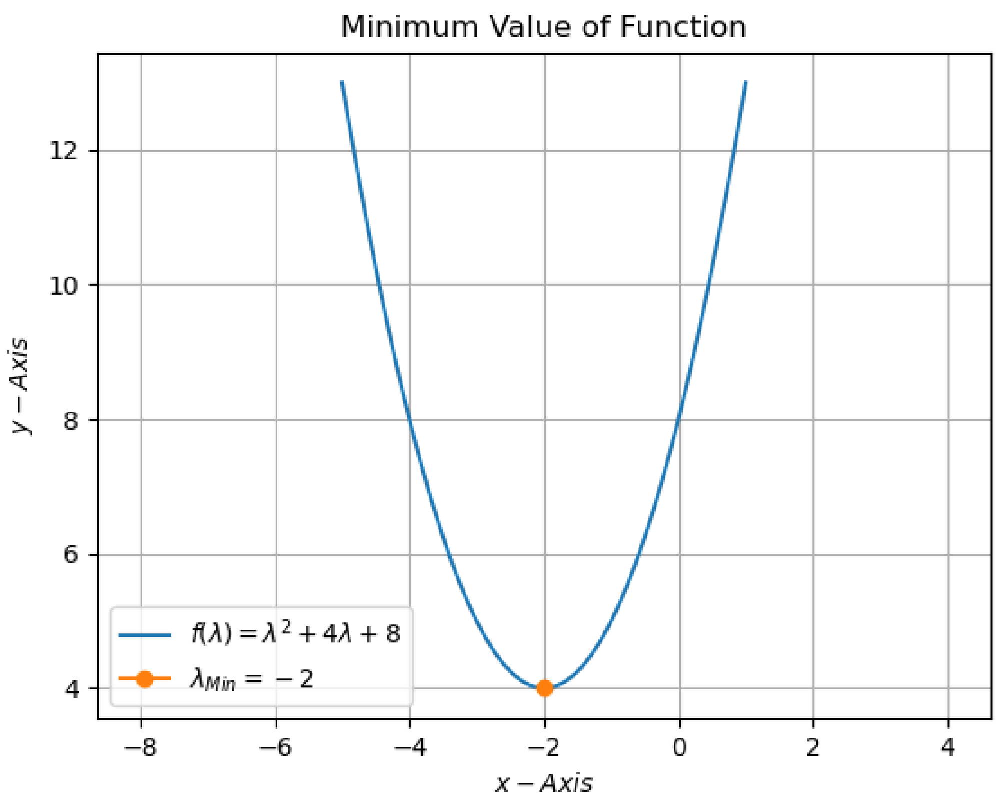


Figure 1.2.2:

convex function The update equation using Gradient Descent is

$$\lambda_{n+1} = \lambda_n - \alpha \nabla f(\lambda_n) \tag{1.2.29}$$

$$= (1 - 2\alpha)\lambda_n - 4\alpha \tag{1.2.30}$$

Taking one-sided Z-transform on both sides of (1.2.30),

$$z\Lambda\left(z\right) = \left(1 - 2\alpha\right)\Lambda\left(z\right) - \frac{4\alpha}{1 - z^{-1}} \tag{1.2.31}$$

$$\Lambda\left(z\right) = -\frac{4\alpha z^{-1}}{\left(1 - \left(1 - 2\alpha\right)z^{-1}\right)\left(1 - z^{-1}\right)} \tag{1.2.32}$$

$$= 2\left(\frac{1}{\left(1 - \left(1 - 2\alpha\right)z^{-1}\right)} - \frac{1}{1 - z^{-1}}\right) \tag{1.2.33}$$

$$= 2\sum_{k=0}^{\infty}\left(\left(1 - 2\alpha\right)^{k} - 1\right)z^{-k} \tag{1.2.34}$$

from (1.2.34), the ROC is

$$\left|z\right| > \max\left\{1, \left|1 - 2\alpha\right|\right\} \tag{1.2.35}$$

$$\implies 0 < \left|1 - 2\alpha\right| < 1 \tag{1.2.36}$$

$$\implies 0 < \alpha < \frac{1}{2} \tag{1.2.37}$$

Thus, if α satisfies (1.2.37), then from (1.2.34)

$$\lim_{n\to\infty}\lambda_n = -2 \tag{1.2.38}$$

Choosing

i. $\alpha = 0.001$

ii. precision = 0.0000001

iii. n = 10000000

iv. $\lambda_0 = 4$

$$\lambda_{min} = -2 \tag{1.2.39}$$

$$\mathbf{x}_{min} = \mathbf{P} = \begin{pmatrix} 2 \\ 2 \end{pmatrix} + (-2) \begin{pmatrix} 1 \\ 0 \end{pmatrix} \tag{1.2.40}$$

$$= \begin{pmatrix} 0 \\ 2 \end{pmatrix} \tag{1.2.41}$$

$$OP = \|\mathbf{P} - \mathbf{O}\| \tag{1.2.42}$$

$$= \left\| \begin{pmatrix} 0 \\ 2 \end{pmatrix} - \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right\| \tag{1.2.43}$$

$$= 2 \tag{1.2.44}$$

See Fig. 1.2.3 and Fig. 1.2.4

(c) Lagrange Multipliers

The given problem can be formulated as

$$\min_{\mathbf{x}} f(\mathbf{x}) = \|\mathbf{x} - \mathbf{O}\|^2 \tag{1.2.45}$$

$$\text{s.t. } g(\mathbf{x}) = \mathbf{n}^\top \mathbf{x} - c = 0 \tag{1.2.46}$$

where

$$\mathbf{n} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \mathbf{O} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, c = 2 \tag{1.2.47}$$

Define

$$H(\mathbf{x}, \lambda) = f(\mathbf{x}) - \lambda g(\mathbf{x}) \tag{1.2.48}$$

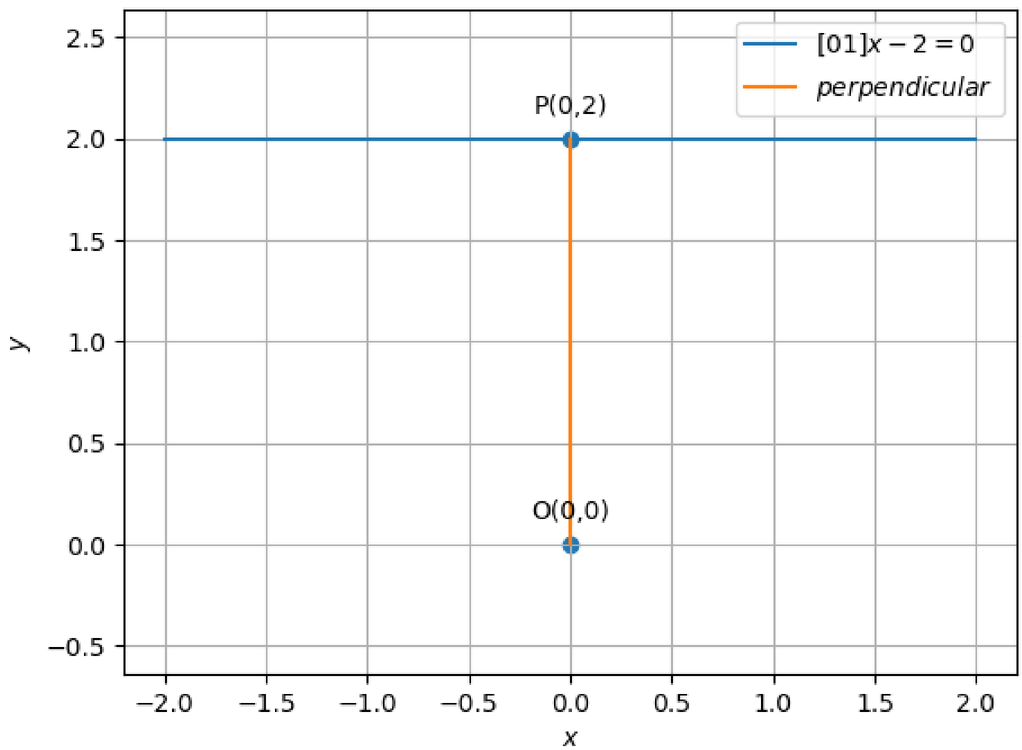


Figure 1.2.3:

Since

$$\nabla f(\mathbf{x}) = 2(\mathbf{x} - \mathbf{O}) \tag{1.2.49}$$

$$\nabla g(\mathbf{x}) = \mathbf{n} \tag{1.2.50}$$

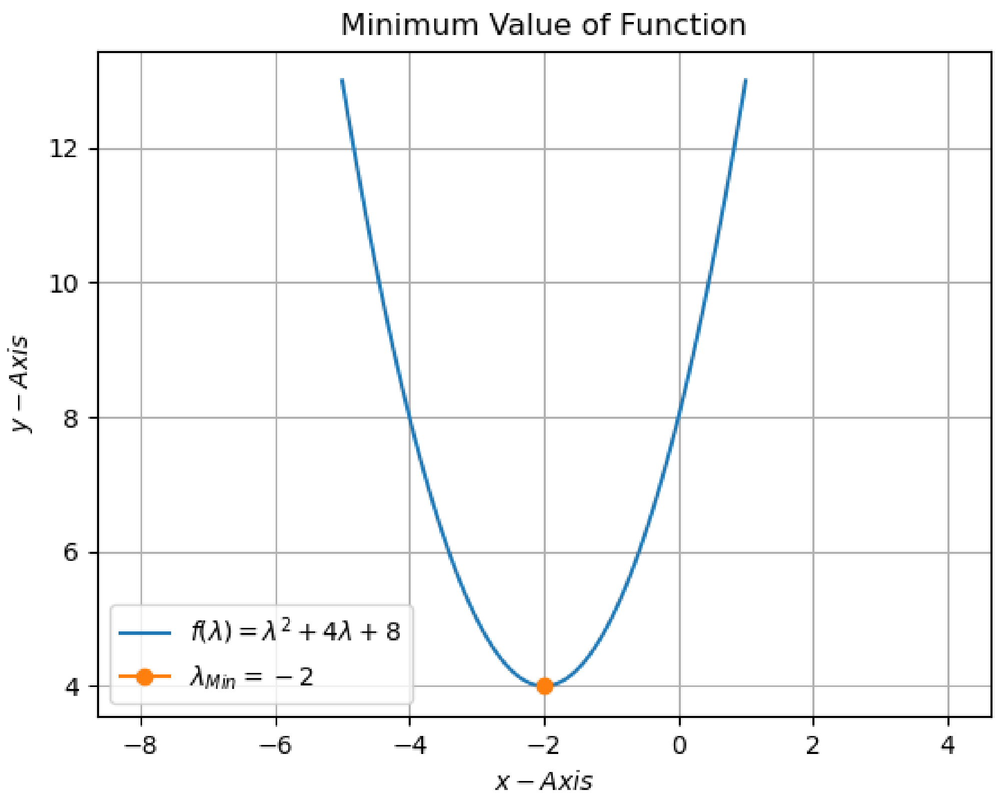


Figure 1.2.4:

We have to find $\lambda \in \mathbb{R}$ such that

$$\nabla H(\mathbf{x}, \lambda) = 0 \tag{1.2.51}$$

$$\implies 2(\mathbf{x} - \mathbf{O}) - \lambda \mathbf{n} = 0 \tag{1.2.52}$$

$$\implies \mathbf{x} = \frac{\lambda}{2}\mathbf{n} + \mathbf{O} \tag{1.2.53}$$

Substituting (1.2.53) in (1.2.46)

$$\mathbf{n}^\top\left(\frac{\lambda}{2}\mathbf{n}+\mathbf{O}\right)-c=0 \tag{1.2.54}$$

$$\implies \lambda=\frac{2\left(c-\mathbf{n}^\top\mathbf{O}\right)}{\|\mathbf{n}\|^2}=4>0 \tag{1.2.55}$$

Substituting the value of λ in (1.2.53),

$$\mathbf{x}_{min}=\mathbf{O}+\frac{\mathbf{n}\left(c-\mathbf{n}^\top\mathbf{O}\right)}{\|\mathbf{n}\|^2} \tag{1.2.56}$$

$$=\begin{pmatrix}0\\0\end{pmatrix}+\frac{\begin{pmatrix}0\\1\end{pmatrix}\left(2-\begin{pmatrix}0&1\end{pmatrix}\begin{pmatrix}0\\0\end{pmatrix}\right)}{1}=\begin{pmatrix}0\\2\end{pmatrix} \tag{1.2.57}$$

$$\implies OP=\|\mathbf{P}-\mathbf{O}\|^2=2 \tag{1.2.58}$$

See Fig. 1.2.5

1.3 Find the perpendicular distance from the origin to the line x–$y=4$ and angle between perpendicular and the positive x-axis.

Solution:

(a) Optimization Problem

The given problem can be expressed as

$$\min_{\mathbf{x}} g\left(\mathbf{x}\right)=\|\mathbf{x}-\mathbf{P}\|^2 \tag{1.3.1}$$

$$\text{s.t.}\quad \mathbf{n}^T\mathbf{x}=c \tag{1.3.2}$$

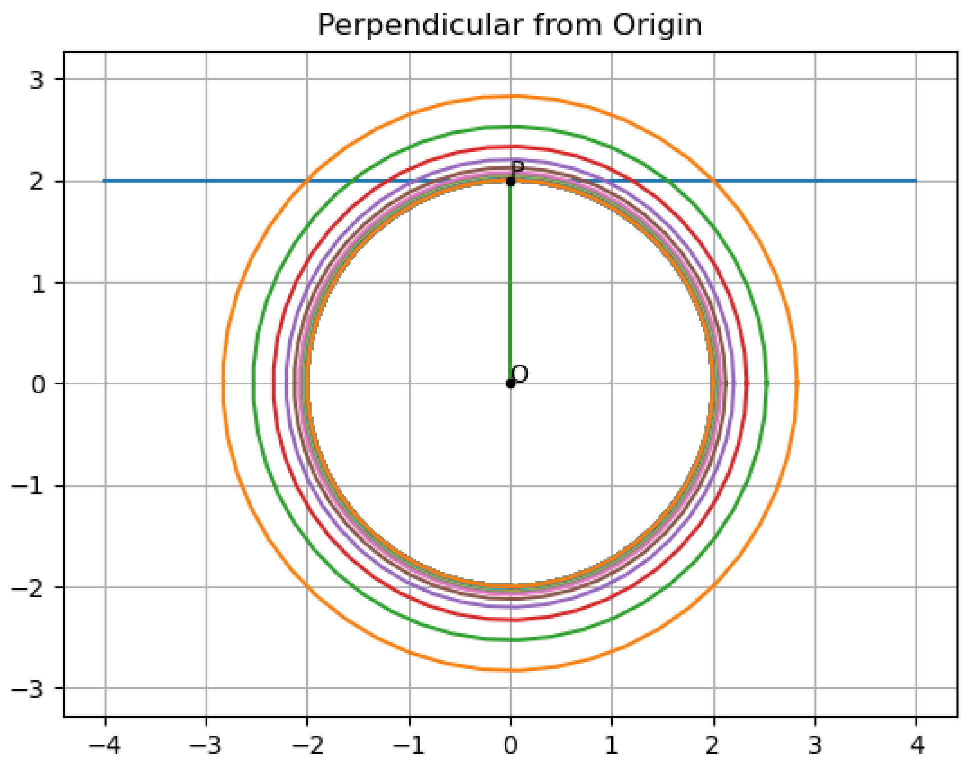


Figure 1.2.5:

where

$$\mathbf{P} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \tag{1.3.3}$$

$$\mathbf{n} = \begin{pmatrix} 1 \\ -1 \end{pmatrix} \tag{1.3.4}$$

$$c = 4 \tag{1.3.5}$$

Solving the equation (1.3.1) using cvxpy we get the solution as,

$$\mathbf{x} = \begin{pmatrix} 2 \\ -2 \end{pmatrix} \tag{1.3.6}$$

The direction vector of the perpendicular is given by,

$$\mathbf{m} = \begin{pmatrix} 1 \\ -1 \end{pmatrix} \tag{1.3.7}$$

The angle between the perpendicular and the positive x-axis is given by,

$$\mathbf{e_1} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \tag{1.3.8}$$

$$\cos\theta = \frac{\mathbf{m}^\top \mathbf{e_1}}{\|\mathbf{m}\| \|\mathbf{e_1}\|} \tag{1.3.9}$$

$$= \frac{1}{\sqrt{2}} \tag{1.3.10}$$

$$\implies \theta = 45 \tag{1.3.11}$$

(b) Gradient Descent

The given problem can be expressed as a constrained optimization problem as

$$\min_{\mathbf{x}} g(\mathbf{x}) = \|\mathbf{x} - \mathbf{P}\|^2 \tag{1.3.12}$$

$$\text{s.t.} \quad \mathbf{n}^T \mathbf{x} = c \tag{1.3.13}$$

where

$$\mathbf{P} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \tag{1.3.14}$$

$$\mathbf{n} = \begin{pmatrix} 1 \\ -1 \end{pmatrix} \tag{1.3.15}$$

$$c = 4 \tag{1.3.16}$$

The problem (1.3.12) can be modified into an unconstrained optimized problem as follows

$$\min_{\lambda} f(\lambda) = \left\|\mathbf{A} + \lambda\mathbf{m} - \mathbf{P}\right\|^2 \tag{1.3.17}$$

where $\mathbf{m}$ is the direction vector of the given line and $\mathbf{A}$ is any point on the line.

$$\mathbf{m} = \begin{pmatrix} -1 \\ 1 \end{pmatrix} \tag{1.3.18}$$

$$\mathbf{A} = \begin{pmatrix} 4 \\ 0 \end{pmatrix} \tag{1.3.19}$$

$$f(\lambda) = \left(\lambda\mathbf{m} + \mathbf{A} - \mathbf{P}\right)^T \left(\lambda\mathbf{m} + \mathbf{A} - \mathbf{P}\right) \tag{1.3.20}$$

$$= \lambda^2 \left\|\mathbf{m}\right\|^2 + 2\lambda\mathbf{m}^T \left(\mathbf{A} - \mathbf{P}\right) + \left\|\mathbf{A} - \mathbf{P}\right\|^2 \tag{1.3.21}$$

Here we have,

$$f(\lambda) = 2\lambda^2 - 8\lambda + 16 \tag{1.3.22}$$

A numerical solution for (1.3.17) is obtained as

$$\lambda_{n+1} = \lambda_n - \alpha f'(\lambda_n) \tag{1.3.23}$$

where λ_0 is an inital guess and μ is a variable parameter. These parameters decide how fast the algorithm converges.

From (1.3.22) we get

$$\lambda_{n+1} = \lambda_n - \alpha(4\lambda_n - 8) \tag{1.3.24}$$

By taking the parameters as listed in the below table

Parameter	**Description**	**Value**
λ_0	Initial guess	-1
α	Variable parameter	0.01
N	Number of iterations	10000
ϵ	Tolerance in λ	10^{-6}

Table 1.3.1:

λ obtained is

$$\lambda = 2 \tag{1.3.25}$$

Hence, from equation (1.3.22)

$$f(2) = 8 \tag{1.3.26}$$

Hence the perpendicular distance is $2\sqrt{2}$

(c) Lagrange Multipliers

The given problem can be expressed as a constrained optimization problem as

$$\min_{\mathbf{x}} f(\mathbf{x}) \triangleq \|\mathbf{x} - \mathbf{P}\|^2 \tag{1.3.27}$$

$$\text{s.t.} \quad g(\mathbf{x}) \triangleq \mathbf{n}^T \mathbf{x} - c = 0 \tag{1.3.28}$$

where

$$\mathbf{P} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \mathbf{n} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}, c = 4 \tag{1.3.29}$$

Define

$$L(\mathbf{x}, \lambda) = f(\mathbf{x}) - \lambda g(\mathbf{x}) \tag{1.3.30}$$

Here we find the optimal point, $\mathbf{Q}$ that is closest to the point $\mathbf{P}$, by finding λ using the following equation

$$\nabla L(\mathbf{x}, \lambda) = 0 \tag{1.3.31}$$

Here we have,

$$\nabla f\left(\mathbf{x}\right) = 2\left(\mathbf{x} - \mathbf{P}\right) \tag{1.3.32}$$

$$\nabla g\left(\mathbf{x}\right) = \mathbf{n} \tag{1.3.33}$$

From (1.3.31), (1.3.32), (1.3.33) we get

$$2\left(\mathbf{x} - \mathbf{P}\right) - \lambda\mathbf{n} = 0 \tag{1.3.34}$$

$$\implies \mathbf{x} = \frac{\lambda}{2}\mathbf{n} + \mathbf{P} \tag{1.3.35}$$

The point $\mathbf{x}$ lies on the given line (1.3.28)

$$\mathbf{n}^\top\left(\frac{\lambda}{2}\mathbf{n} + \mathbf{P}\right) - c = 0 \tag{1.3.36}$$

$$\implies \lambda = -\frac{2\left(\mathbf{n}^\top\mathbf{P} - c\right)}{\left\|\mathbf{n}\right\|^2} \tag{1.3.37}$$

Substituting (1.3.37) in (1.3.35), we get the optimal point $\mathbf{Q}$ as

$$\mathbf{Q} = \mathbf{P} - \frac{\mathbf{n}^\top\mathbf{P} - c}{\left\|\mathbf{n}\right\|^2}\mathbf{n} \tag{1.3.38}$$

Substituting (1.3.29) in (1.3.37), (1.3.38) gives

$$\lambda = 4 \tag{1.3.39}$$

$$\mathbf{Q} = \begin{pmatrix} 2 \\ -2 \end{pmatrix} \tag{1.3.40}$$

Hence the perpendicular distance is given by

$$d = \|\mathbf{Q} - \mathbf{P}\| \tag{1.3.41}$$

$$= 2\sqrt{2} \tag{1.3.42}$$

1.4 The line through the points $(h, 3)$ and $(4, 1)$ intersects the line $7x - 9y - 19 = 0$ at right angle. Find the value of h and the intersection of the lines.

Solution:

(a) Optimization Problem

Let the point $\mathbf{P}$ be the foot of the perpendicular on the line $7x-9y-19 = 0$ from point $\begin{pmatrix} h \\ 3 \end{pmatrix}$ (Let's say point $\mathbf{O}$). The optimization problem can be expressed as

$$\min_{\mathbf{x}} \|\mathbf{x} - \mathbf{O}\|^2 \tag{1.4.1}$$

$$\text{s.t.} \quad \mathbf{n}^\top \mathbf{x} = c \tag{1.4.2}$$

where

$$\mathbf{n} = \begin{pmatrix} 7 \\ -9 \end{pmatrix}, c = 19 \tag{1.4.3}$$

The line equation can be expressed as

$$\mathbf{x} = \mathbf{A} + \lambda \mathbf{m} \tag{1.4.4}$$

where

$$\mathbf{m} = \begin{pmatrix} 9 \\ 7 \end{pmatrix}, \mathbf{A} = \begin{pmatrix} \frac{19}{7} \\ 0 \end{pmatrix} \tag{1.4.5}$$

Using the parameric form, Substituting (1.4.4) in (1.4.1), the optimization problem becomes

$$\min_{\lambda} \left\| \lambda \mathbf{m} + (\mathbf{A} - \mathbf{O}) \right\|^2 \tag{1.4.6}$$

$$\begin{gathered} \implies \min_{\lambda} f(\lambda) = \lambda^2 \left\| \mathbf{m} \right\|^2 + \\ 2\lambda \left(\mathbf{A} - \mathbf{O} \right)^\top \mathbf{m} + \left\| \mathbf{A} - \mathbf{O} \right\|^2 \end{gathered} \tag{1.4.7}$$

$\because$ the coefficient of $\lambda^2 > 0$, (1.4.7) is a convex function. Thus,

$$f''(\lambda) = 2 \left\| \mathbf{m} \right\|^2 \tag{1.4.8}$$

$$\because f''(\lambda) > 0, f'(\lambda_{min}) = 0, \text{ for } \lambda_{min} \tag{1.4.9}$$

yielding

$$f'(\lambda_{min}) = 2\lambda_{min} \left\| \mathbf{m} \right\|^2 + 2 \left(\mathbf{A} - \mathbf{O} \right)^\top \mathbf{m} = 0 \tag{1.4.10}$$

$$\lambda_{min} = -\frac{\left(\mathbf{A} - \mathbf{O} \right)^\top \mathbf{m}}{\left\| \mathbf{m} \right\|^2} \tag{1.4.11}$$

It is given that the line through the points $\begin{pmatrix} h \\ 3 \end{pmatrix}$ and $\begin{pmatrix} 4 \\ 1 \end{pmatrix}$ intersects the line

$7x - 9y - 19 = 0$ at right angle. And the point $\begin{pmatrix} 4 \\ 1 \end{pmatrix}$ is on the line $7x - 9y - 19 = 0$.

From equation (1.4.4)

$$\implies \begin{pmatrix} 4 \\ 1 \end{pmatrix} = \begin{pmatrix} \frac{19}{7} \\ 0 \end{pmatrix} + \lambda_{min} \begin{pmatrix} 9 \\ 7 \end{pmatrix} \tag{1.4.12}$$

$$\implies \lambda_{min} = \frac{1}{7} \tag{1.4.13}$$

Substituting the values of $\mathbf{A}$, $\mathbf{O}$, λ_{min} and $\mathbf{m}$ in equation (1.4.11)

$$\frac{1}{7} = -\frac{\left(\begin{pmatrix} \frac{19}{7} \\ 0 \end{pmatrix} - \begin{pmatrix} h \\ 3 \end{pmatrix} \right)^\top \begin{pmatrix} 9 \\ 7 \end{pmatrix}}{\left\| \begin{pmatrix} 9 \\ 7 \end{pmatrix} \right\|^2} \tag{1.4.14}$$

$$\implies \frac{130}{7} = -\frac{171}{7} + 9h + 21 \tag{1.4.15}$$

$$\implies h = \frac{22}{9} \tag{1.4.16}$$

The relavent figure is 1.4.1

1.5 Find the coordinates of the foot of perpendicular from the point

$$\mathbf{P} = \begin{pmatrix} -1 \\ 3 \end{pmatrix} \tag{1.5.1}$$

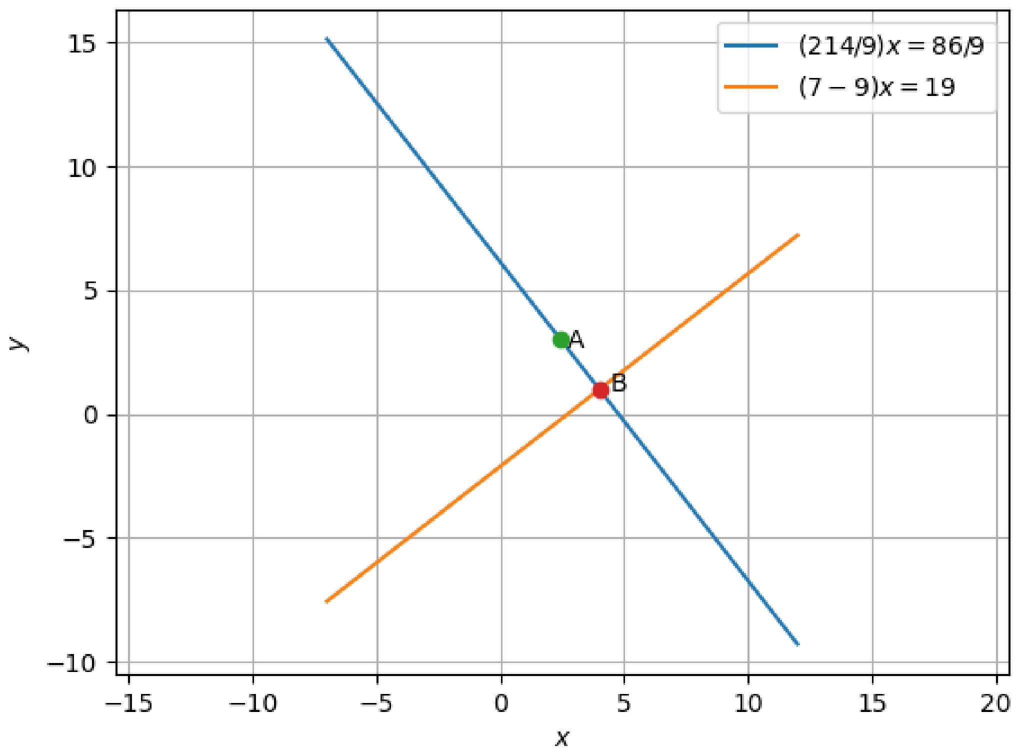


Figure 1.4.1:

to the line

$$\begin{pmatrix} 3 & -4 \end{pmatrix} \mathbf{x} = 16 \tag{1.5.2}$$

Solution:

(a) Optimization Problem

Any point on (1.5.2) is clearly of the form

$$\mathbf{Q} = \mathbf{A} + \lambda \mathbf{m} \tag{1.5.3}$$

where $\lambda \in \mathbb{R}$ and

$$\mathbf{A} = \begin{pmatrix} 0 \\ -4 \end{pmatrix}, \ \mathbf{m} = \begin{pmatrix} 4 \\ 3 \end{pmatrix} \tag{1.5.4}$$

Thus,

$$f(\lambda) = \left\|\mathbf{Q} - \mathbf{P}\right\|^2 \tag{1.5.5}$$

$$= \left\|\mathbf{A} - \mathbf{P} + \lambda\mathbf{m}\right\|^2 \tag{1.5.6}$$

$$= \left\|\mathbf{m}\right\|^2 \lambda^2 + 2\mathbf{m}^\top \left(\mathbf{A} - \mathbf{P}\right) \lambda + \left\|\mathbf{A} - \mathbf{P}\right\|^2 \tag{1.5.7}$$

Since the coefficient of λ^2 in $f(\lambda)$ is positive, it follows that $f(\lambda)$ is convex. Hence, the minima is achieved at

$$f'(\lambda_m) = 2\left(\left\|\mathbf{m}\right\|^2 \lambda_m + \mathbf{m}^\top \left(\mathbf{A} - \mathbf{P}\right)\right) = 0 \tag{1.5.8}$$

$$\implies \lambda_m = -\frac{\mathbf{m}^\top \left(\mathbf{A} - \mathbf{P}\right)}{\left\|\mathbf{m}\right\|^2} \tag{1.5.9}$$

Thus,

$$\mathbf{Q}_{\mathbf{m}} = \mathbf{A} + \lambda_m \mathbf{m} \tag{1.5.10}$$

$$= \mathbf{A} - \frac{\mathbf{m}^\top \left(\mathbf{A} - \mathbf{P}\right)}{\left\|\mathbf{m}\right\|^2}\mathbf{m} \tag{1.5.11}$$

Thus, substituting (1.5.4) into (1.5.11), we get

$$\mathbf{Q_m} = \frac{1}{25}\begin{pmatrix} 68 \\ -49 \end{pmatrix} \tag{1.5.12}$$

The value of λ_m is verified in Fig. 1.5.1.

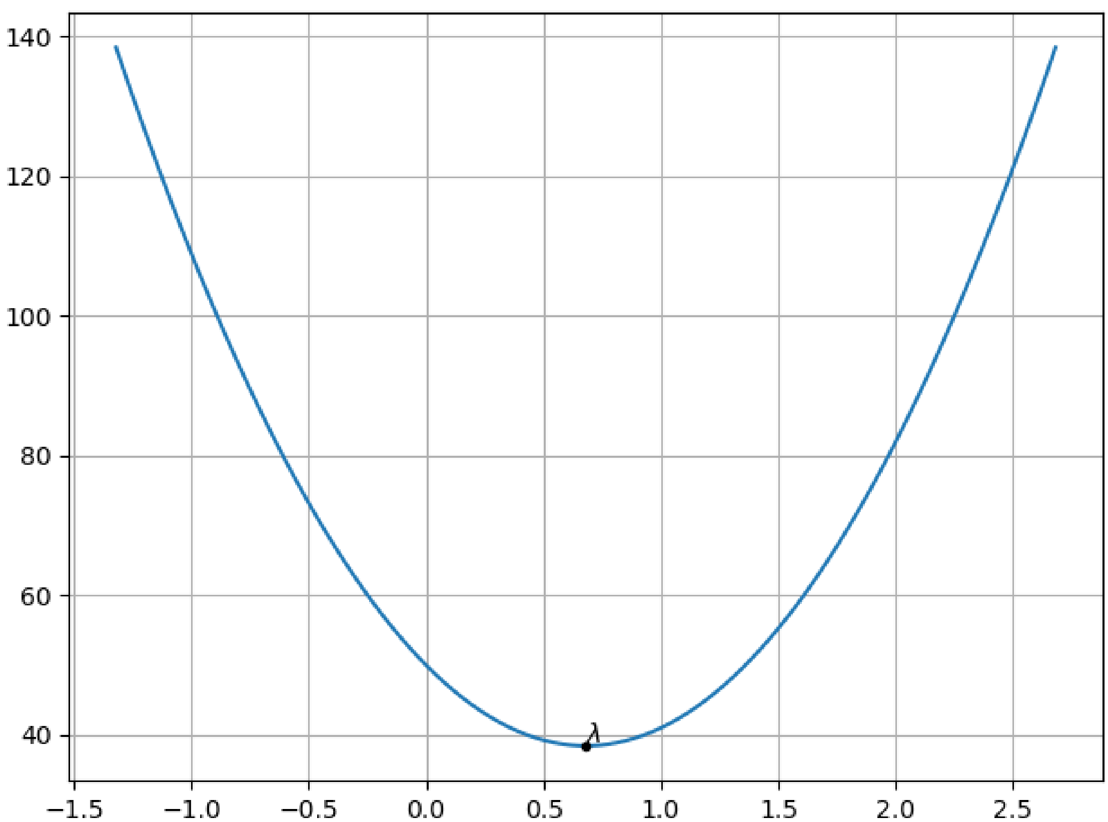


Figure 1.5.1: This convex function achieves its minimum at λ_m.

(b) Gradient Descent

Any point on (1.5.2) is clearly of the form

$$\mathbf{Q} = \mathbf{A} + \lambda \mathbf{m} \tag{1.5.13}$$

where $\lambda \in \mathbb{R}$ and

$$\mathbf{A} = \begin{pmatrix} 0 \\ -4 \end{pmatrix}, \ \mathbf{m} = \begin{pmatrix} 4 \\ 3 \end{pmatrix} \tag{1.5.14}$$

Thus,

$$f(\lambda) = \|\mathbf{Q} - \mathbf{P}\|^2 \tag{1.5.15}$$

$$= \|\mathbf{A} - \mathbf{P} + \lambda \mathbf{m}\|^2 \tag{1.5.16}$$

$$= \|\mathbf{m}\|^2 \lambda^2 + 2\mathbf{m}^\top (\mathbf{A} - \mathbf{P}) \lambda + \|\mathbf{A} - \mathbf{P}\|^2 \tag{1.5.17}$$

Since (1.5.17) is convex, we use the gradient descent function on λ to converge at the minimum of $f(\lambda)$.

$$\lambda_{n+1} = \lambda_n - \alpha f'(\lambda_n) \tag{1.5.18}$$

$$= \left(1 - 2\alpha \|\mathbf{m}\|^2\right) \lambda_n + 2\alpha \mathbf{m}^\top (\mathbf{A} - \mathbf{P}) \tag{1.5.19}$$

Taking the one-sided Z-transform on both sides of (1.5.19),

$$z\Lambda(z) = \left(1 - 2\alpha \|\mathbf{m}\|^2\right) \Lambda(z) - \frac{2\alpha \mathbf{m}^\top (\mathbf{A} - \mathbf{P})}{1 - z^{-1}} \tag{1.5.20}$$

Solving (1.5.20)

$$\Lambda(z) = -\frac{2\alpha \mathbf{m}^\top (\mathbf{A} - \mathbf{P}) z^{-1}}{(1 - z^{-1}) \left(1 - \left(1 - 2\alpha \|\mathbf{m}\|^2\right) z^{-1}\right)} \tag{1.5.21}$$

$$= -\frac{\mathbf{m}^\top (\mathbf{A} - \mathbf{P})}{\|\mathbf{m}\|^2} \left(\frac{1}{1 - z^{-1}}\right. \tag{1.5.22}$$

$$\left. - \frac{1}{1 - \left(1 - 2\alpha \|\mathbf{m}\|^2\right) z^{-1}}\right) \tag{1.5.23}$$

$$= \frac{\mathbf{m}^\top (\mathbf{A} - \mathbf{P})}{\|\mathbf{m}\|^2} \sum_{k=0}^{\infty} \left(1 - \left(1 - 2\alpha \|\mathbf{m}\|^2\right)^k\right) z^{-k} \tag{1.5.24}$$

From (1.5.21), the ROC is

$$|z| > \max\left\{1, 1 - 2\alpha \|\mathbf{m}\|^2\right\} \tag{1.5.25}$$

$$\implies 0 < 1 - 2\alpha \|\mathbf{m}\|^2 < 1 \tag{1.5.26}$$

$$\implies 0 < \alpha < \frac{1}{2 \|\mathbf{m}\|^2} \tag{1.5.27}$$

Thus, if α satisfies (1.5.27), then from (1.5.24), substituting from (1.5.14),

$$\lim_{n \to \infty} \lambda_n = -\frac{\mathbf{m}^\top (\mathbf{A} - \mathbf{P})}{\|\mathbf{m}^2\|} = \frac{17}{25} \tag{1.5.28}$$

We select the following parameters to arrive at the optimal λ, where N is the number of iterations and ϵ is the convergence limit. The gradient descent is demonstrated in Fig. 1.5.2, plotted by the Python code. The relevant parameters are shown in Table 1.5.1.

(c) Lagrange Multipliers

Parameter	Value
λ_0	0
α	0.1
N	1000000
ϵ	10^{-6}

Table 1.5.1: Parameters for Gradient Descent

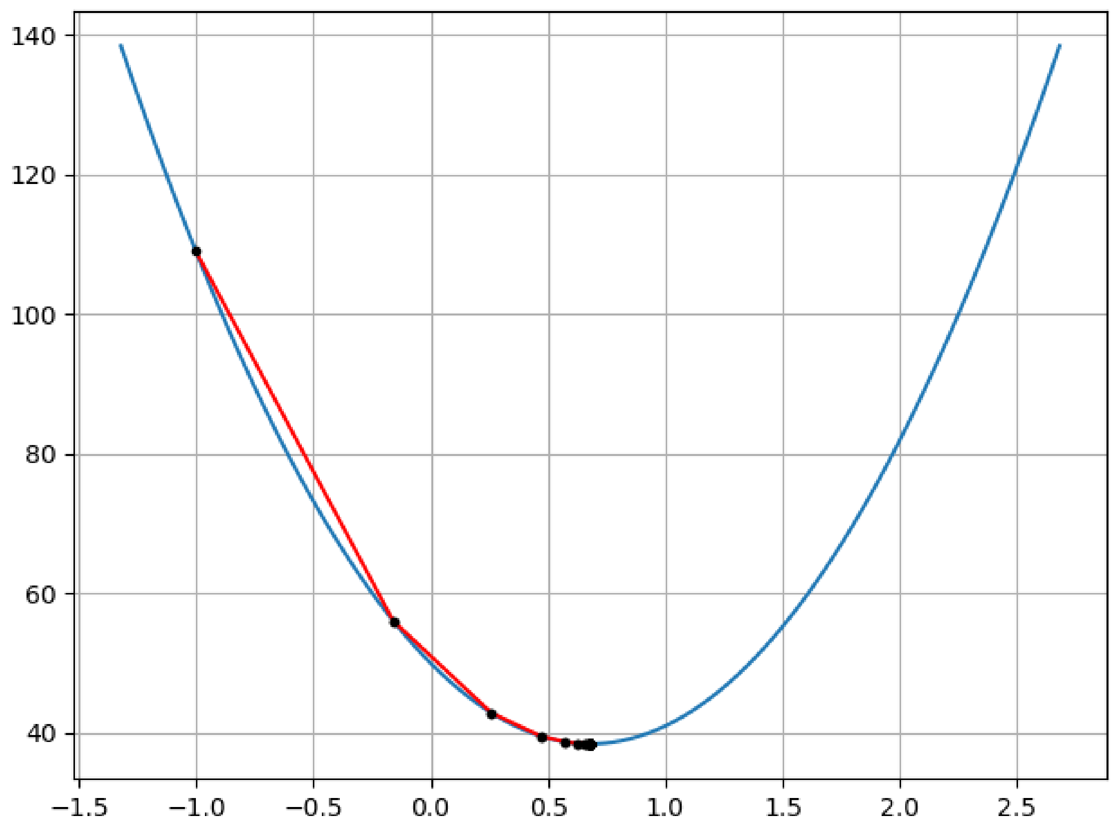


Figure 1.5.2: Gradient descent to get the optimal λ.

Solution: We rewrite the problem as

$$\min_{\mathbf{x}} h(\mathbf{x}) \triangleq \|\mathbf{x} - \mathbf{P}\|^2 \tag{1.5.29}$$

$$\text{s.t. } g(\mathbf{x}) \triangleq \mathbf{n}^\top \mathbf{x} - c = 0 \tag{1.5.30}$$

where

$$\mathbf{P} = \begin{pmatrix} -1 \\ 3 \end{pmatrix}, \ \mathbf{n} = \begin{pmatrix} 3 \\ -4 \end{pmatrix}, \ c = 16 \tag{1.5.31}$$

Define

$$C(\mathbf{x}, \lambda) = h(\mathbf{x}) - \lambda g(\mathbf{x}) \tag{1.5.32}$$

and note that

$$\nabla h(\mathbf{x}) = 2(\mathbf{x} - \mathbf{P}) \tag{1.5.33}$$

$$\nabla g(\mathbf{x}) = \mathbf{n} \tag{1.5.34}$$

We are required to find $\lambda \in \mathbb{R}$ such that

$$\nabla C(\mathbf{x}, \lambda) = 0 \tag{1.5.35}$$

$$\implies 2(\mathbf{x} - \mathbf{P}) - \lambda \mathbf{n} = 0 \tag{1.5.36}$$

However, $\mathbf{x}$ lies on the line (1.5.2). Thus, from (1.5.36),

$$\mathbf{n}^\top \left(\frac{\lambda}{2} \mathbf{n} + \mathbf{P} \right) - c = 0 \tag{1.5.37}$$

$$\implies \lambda = \frac{2\left(c - \mathbf{n}^\top \mathbf{P}\right)}{\|\mathbf{n}\|^2} \tag{1.5.38}$$

Substituting (1.5.38) in (1.5.36), the optimal point is given by

$$\mathbf{Q} = \mathbf{P} + \frac{\lambda}{2}\mathbf{n} \tag{1.5.39}$$

$$= \mathbf{P} - \frac{\mathbf{n}^\top \mathbf{P} - c}{\left\|\mathbf{n}\right\|^2}\mathbf{n} \tag{1.5.40}$$

Substituting from (1.5.31),

$$\lambda = \frac{62}{25}, \ \mathbf{Q} = \frac{1}{25}\begin{pmatrix} 68 \\ -49 \end{pmatrix} \tag{1.5.41}$$

To find $\mathbf{Q}$ graphically, we use constrained gradient descent, with learning rate $\alpha = 0.01$. The results are shown in Fig. 1.5.3, plotted using the Python code. *Constrained gradient descent* is a method of optimizing the cost function subject to some constraints, represented as follows.

$$\max_{\mathbf{x}} f\left(\mathbf{x}\right) \tag{1.5.42}$$

$$\text{s.t. } g\left(\mathbf{x}\right) = 0 \tag{1.5.43}$$

Unlike the unconstrained version, one cannot move in the negative direction of the gradient vector of $f\left(\mathbf{x}\right)$. However, we must move along the constraint in (1.5.43).

The algorithm terminates when the gradient vector of f is parallel to the normal vector of g at that point. Mathematically, at an optimum $\mathbf{x_o}$,

$$\nabla f\left(\mathbf{x_o}\right) = \lambda \nabla g\left(\mathbf{x_o}\right) \tag{1.5.44}$$

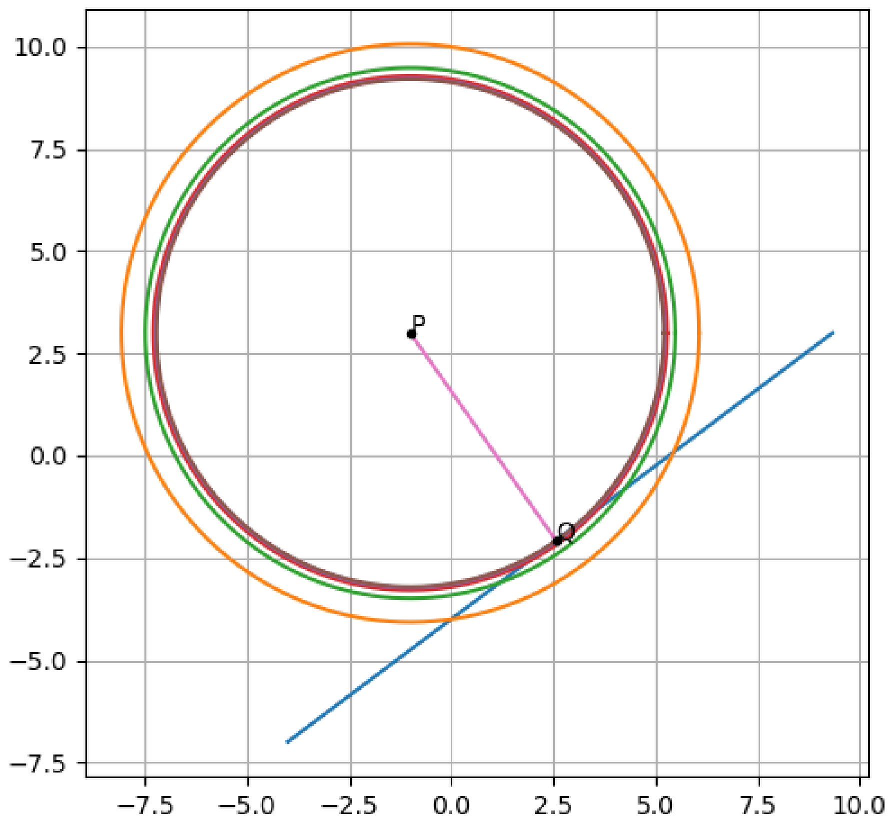


Figure 1.5.3: Constrained gradient descent to find optimal $\mathbf{Q}$.

where $\lambda \in \mathbb{R} \setminus \{0\}$. Observe that (1.5.44) may be rewritten as

$$\nabla C\left(\mathbf{x}, \lambda\right) = \nabla\left(f\left(\mathbf{x}\right) - \lambda g\left(\mathbf{x}\right)\right) = 0 \tag{1.5.45}$$

which is analogous to the method of Lagrangian multipliers.

1.6 Find the shortest distance between the lines l_1 and l_2 whose vector equations are $\overrightarrow{r} = \hat{i} + \hat{j} + \lambda(2\hat{i} - \hat{j} + \hat{k})$ and $\overrightarrow{r} = 2\hat{i} + \hat{j} - \hat{k} + \mu(3\hat{i} - 5\hat{j} + 2\hat{k})$

Solution:

(a) Optimization Problem

The lines l_1 and l_2 in vector form can be written as

$$\mathbf{x} = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + \lambda_1 \begin{pmatrix} 2 \\ -1 \\ 1 \end{pmatrix} \tag{1.6.1}$$

$$\mathbf{x} = \begin{pmatrix} 2 \\ 1 \\ -1 \end{pmatrix} + \lambda_2 \begin{pmatrix} 3 \\ -5 \\ 2 \end{pmatrix} \tag{1.6.2}$$

$$\mathbf{x_1} = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \mathbf{x_2} = \begin{pmatrix} 2 \\ 1 \\ -1 \end{pmatrix}, \mathbf{m_1} = \begin{pmatrix} 2 \\ -1 \\ 1 \end{pmatrix}, \mathbf{m_2} = \begin{pmatrix} 3 \\ -5 \\ 2 \end{pmatrix} \tag{1.6.3}$$

The distance between the lines is given by,

$$d = \|(\mathbf{x_2} + \lambda_2 \mathbf{m_2}) - (\mathbf{x_1} + \lambda_1 \mathbf{m_1})\| \tag{1.6.4}$$

$$\implies d = \|\mathbf{x_2} - \mathbf{x_1} - \lambda_1 \mathbf{m_1} + \lambda_2 \mathbf{m_2}\| \tag{1.6.5}$$

Consider the following definitions

$$\mathbf{A} \triangleq \mathbf{x_2} - \mathbf{x_1}, \mathbf{M} \triangleq \begin{pmatrix} \mathbf{m_1} & \mathbf{m_2} \end{pmatrix}, \boldsymbol{\lambda} \triangleq \begin{pmatrix} \lambda_1 \\ -\lambda_2 \end{pmatrix} \tag{1.6.6}$$

From (1.6.6),

$$d = \|\mathbf{A} - \mathbf{M}\boldsymbol{\lambda}\| \tag{1.6.7}$$

Here we have the values of $\mathbf{A}, \mathbf{M}$ as

$$\mathbf{M} = \begin{pmatrix} 2 & 3 \\ -1 & -5 \\ 1 & 2 \end{pmatrix}, \mathbf{A} = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} \tag{1.6.8}$$

The given problem can be formulated as

$$\min_{\boldsymbol{\lambda}} d^2 = \boldsymbol{\lambda}^\top \mathbf{M}^\top \mathbf{M} \boldsymbol{\lambda} - 2\mathbf{A}^\top \mathbf{M} \boldsymbol{\lambda} + \mathbf{A}^\top \mathbf{A} \tag{1.6.9}$$

$$\text{s.t.} \quad \boldsymbol{\lambda} \in \mathbb{R}^2 \tag{1.6.10}$$

By solving using cvxpy, we get

$$\min_{\boldsymbol{\lambda}} d = 1.3019 \tag{1.6.11}$$

$$\boldsymbol{\lambda} = \begin{pmatrix} 0.4237 \\ -0.1186 \end{pmatrix} \tag{1.6.12}$$

The shortest distance between the given lines is 1.3019 units.

(b) Gradient Descent

A numerical solution is obtained as

$$\boldsymbol{\lambda}_{n+1} = \boldsymbol{\lambda}_n - \alpha \nabla f\left(\boldsymbol{\lambda_n}\right) \tag{1.6.13}$$

where λ_0 is an inital guess and α is a variable parameter. These parameters decide how fast the algorithm converges. Here the gradient is given by,

$$\nabla f\left(\boldsymbol{\lambda}\right) = 2\mathbf{M}^\top \mathbf{M} \boldsymbol{\lambda} - 2\left(\mathbf{A}^\top \mathbf{M}\right)^\top \tag{1.6.14}$$

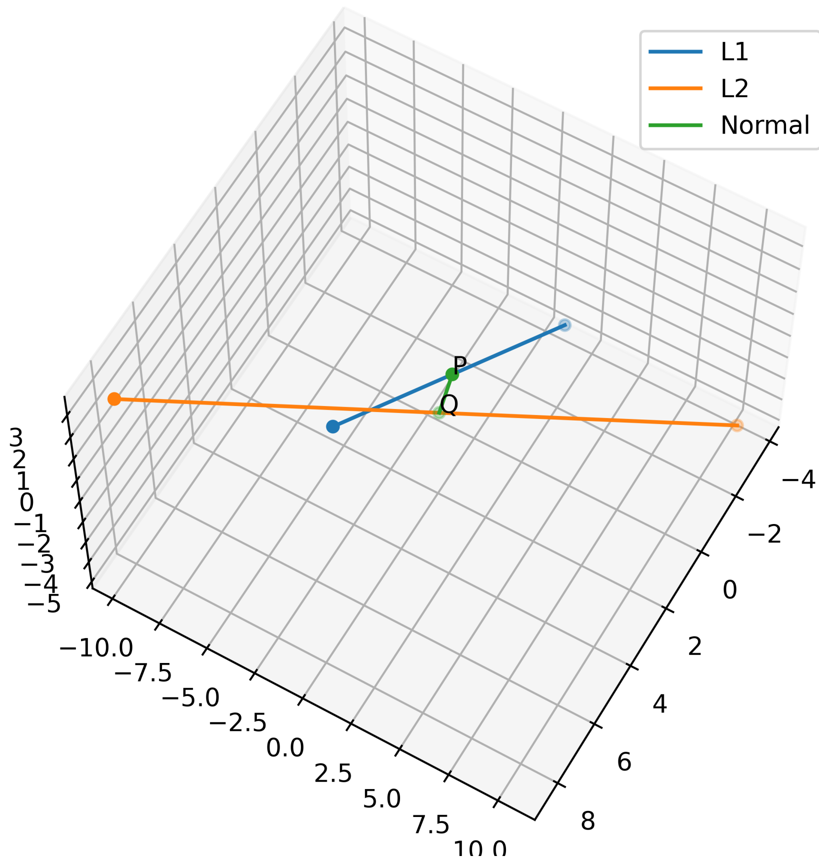


Figure 1.6.1: PQ is the required shortest distance.

From (1.6.13) we get

$$\boldsymbol{\lambda}_{n+1} = \boldsymbol{\lambda}_n - \alpha \left(2\mathbf{M}^\top \mathbf{M} \boldsymbol{\lambda}_n - 2\mathbf{M}^\top \mathbf{A} \right) \tag{1.6.15}$$

By taking the parameters as listed in Table 1.6.1,

$$\boldsymbol{\lambda} = \begin{pmatrix} 0.4237 \\ -0.1186 \end{pmatrix} \tag{1.6.16}$$

Parameter	**Description**	**Value**
λ_0	Initial guess	$\begin{pmatrix}1\\2\end{pmatrix}$
α	Variable parameter	0.01
N	Number of iterations	10000
ϵ	Tolerance in λ	10^{-6}

Table 1.6.1:

The shortest distance is

$$\min_{\boldsymbol{\lambda}} d = 1.3019 \tag{1.6.17}$$

(c) Lagrange Multipliers

Solution: As there are no constraints on $\boldsymbol{\lambda}$, using Lagrange multipliers,

$$\nabla f\left(\boldsymbol{\lambda}\right) = 0 \tag{1.6.18}$$

Here we have the gradient as

$$\nabla f\left(\boldsymbol{\lambda}\right) = 2\mathbf{M}^\top\mathbf{M}\boldsymbol{\lambda} - 2\left(\mathbf{A}^\top\mathbf{M}\right)^\top = 0 \tag{1.6.19}$$

$$\implies \mathbf{M}^\top\mathbf{M}\boldsymbol{\lambda} = \mathbf{M}^\top\mathbf{A} \tag{1.6.20}$$

Solving (1.6.20) we get

$$\boldsymbol{\lambda} = \begin{pmatrix} 0.4237 \\ -0.1186 \end{pmatrix} \tag{1.6.21}$$

The shortest distance is

$$\min_{\boldsymbol{\lambda}} d = 1.3019 \tag{1.6.22}$$

This is also known as the least squares method.

Chapter 2

Distance from a Conic

2.1 The point on the curve

$$x^2 = 2y \tag{2.1.1}$$

which is nearest to the point $\mathbf{P} = \begin{pmatrix} 0 \\ 5 \end{pmatrix}$ is

(a) $\begin{pmatrix} 2\sqrt{2} \\ 4 \end{pmatrix}$

(b) $\begin{pmatrix} 2\sqrt{2} \\ 0 \end{pmatrix}$

(c) $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$

(d) $\begin{pmatrix} 2 \\ 2 \end{pmatrix}$

Solution:

(a) Optimization Problem

(b) Gradient Descent

We need to find

$$\min_{\mathbf{x}} g(\mathbf{x}) = \|\mathbf{x} - \mathbf{P}\|^2 \tag{2.1.2}$$

$$\text{s.t. } h(\mathbf{x}) = \mathbf{x}^\top \mathbf{V}\mathbf{x} + 2\mathbf{u}^\top \mathbf{x} = 0 \tag{2.1.3}$$

where

$$\mathbf{V} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \ \mathbf{u} = \begin{pmatrix} 0 \\ -1 \end{pmatrix} \tag{2.1.4}$$

We find the required minima using constrained gradient descent in Fig. 2.1.1, plotted using the Python code `codes/grad_pits.py`.

(c) Convex Constraint

We need to find

$$\min_{\mathbf{x}} g(\mathbf{x}) = \|\mathbf{x} - \mathbf{P}\|^2 \tag{2.1.5}$$

$$\text{s.t. } h(\mathbf{x}) = \mathbf{x}^\top \mathbf{V}\mathbf{x} + 2\mathbf{u}^\top \mathbf{x} = 0 \tag{2.1.6}$$

where

$$\mathbf{V} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \ \mathbf{u} = \begin{pmatrix} 0 \\ -1 \end{pmatrix} \tag{2.1.7}$$

We find the required minima using constrained gradient descent in Fig. 2.1.2, plotted using Python.

(d) Lagrange Multipliers

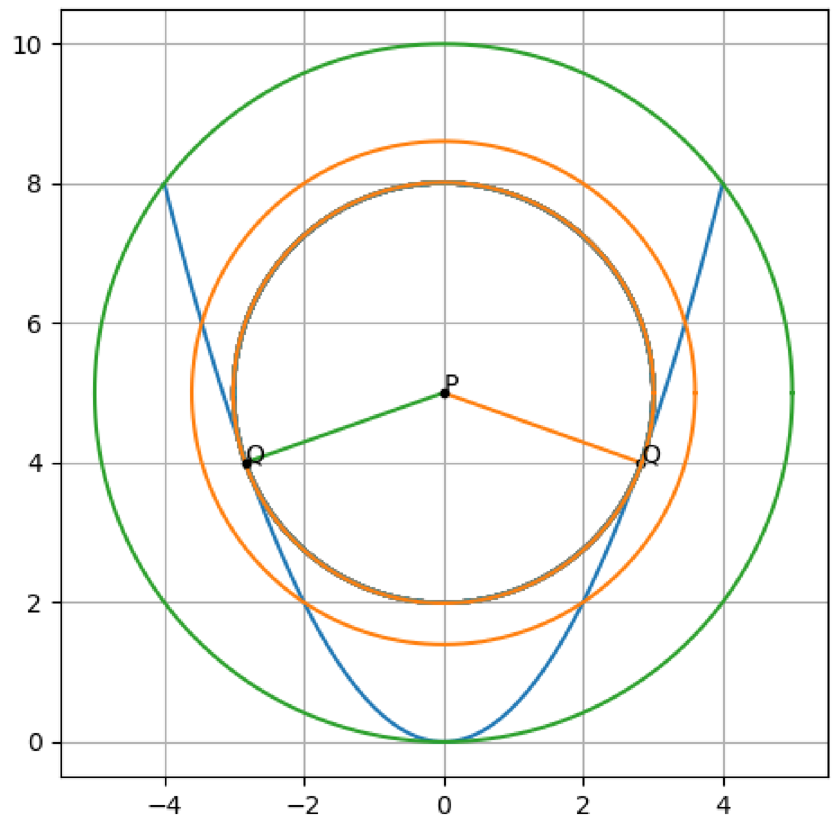


Figure 2.1.1: Gradient descent for a nonconvex optimization problem.

We need to find

$$\min_{\mathbf{x}} g(\mathbf{x}) = \|\mathbf{x} - \mathbf{P}\|^2 \tag{2.1.8}$$

$$\text{s.t. } h(\mathbf{x}) = \mathbf{x}^\top \mathbf{V}\mathbf{x} + 2\mathbf{u}^\top \mathbf{x} = 0 \tag{2.1.9}$$

where

$$\mathbf{V} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \ \mathbf{u} = \begin{pmatrix} 0 \\ -1 \end{pmatrix} \tag{2.1.10}$$

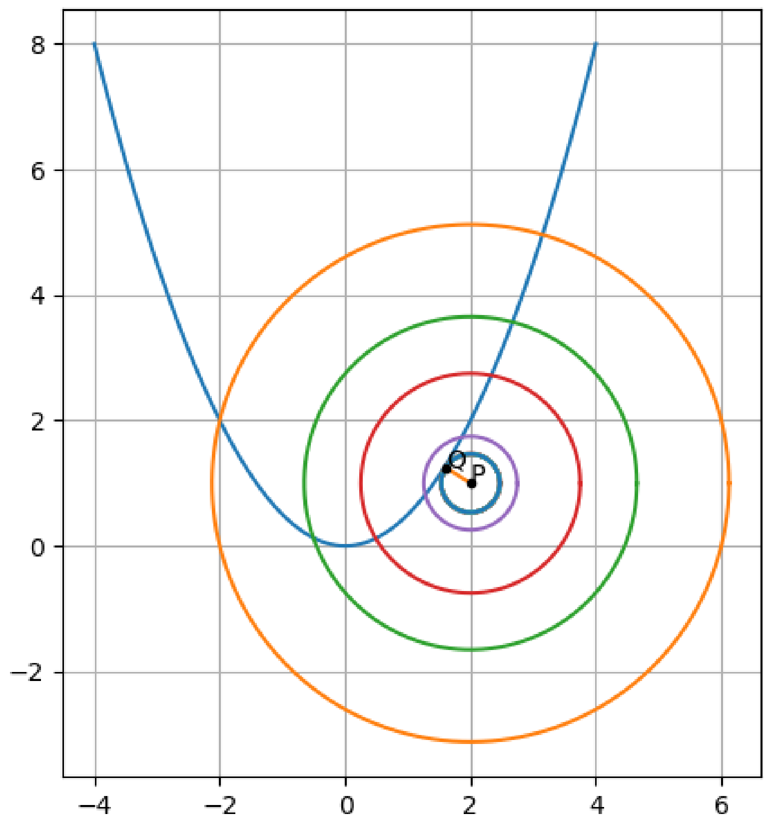


Figure 2.1.2: Gradient descent for a convex optimization problem.

Since the given optimization problem is nonconvex, we use the method of Lagrange multipliers to find the optima. Here, we need to find $\lambda \in \mathbb{R}$ such that there exists a $\mathbf{x}$ satisfying

$$\nabla g(\mathbf{x}) = \lambda \nabla h(\mathbf{x}) \tag{2.1.11}$$

$$\implies 2(\mathbf{x} - \mathbf{P}) = 2\lambda(\mathbf{A}\mathbf{x} + \mathbf{u}) \tag{2.1.12}$$

$$\implies (\mathbf{I} - \lambda\mathbf{A})\mathbf{x} = \lambda\mathbf{u} + \mathbf{P} \tag{2.1.13}$$

$$\implies \begin{pmatrix} 1-\lambda & 0 \\ 0 & 1 \end{pmatrix} \mathbf{x} = \begin{pmatrix} 0 \\ 5-\lambda \end{pmatrix} \tag{2.1.14}$$

From (2.1.14), we have two cases:

i. $\lambda \neq 1$. In this case, we form the augmented matrix

$$\begin{pmatrix} 1-\lambda & 0 & 0 \\ 0 & 1 & 5-\lambda \end{pmatrix} \xleftrightarrow{R_1 \leftarrow \frac{R_1}{1-\lambda}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 5-\lambda \end{pmatrix} \tag{2.1.15}$$

and get that

$$\mathbf{x_m} = \begin{pmatrix} 0 \\ 5-\lambda \end{pmatrix} \tag{2.1.16}$$

Substituting in (2.1.9) gives $\lambda = 5$. Thus, $\mathbf{x_m} = \mathbf{0}$.

ii. $\lambda = 1$. In this case, (2.1.14) becomes

$$\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \mathbf{x} = \begin{pmatrix} 0 \\ 4 \end{pmatrix} \tag{2.1.17}$$

$$\implies \mathbf{e_2}^\top \mathbf{x} = 4 \tag{2.1.18}$$

Substituting (2.1.18) into (2.1.9) becomes

$$\left(\mathbf{e_1}^\top \mathbf{x}\right)^2 = 8 \tag{2.1.19}$$

$$\implies \mathbf{e_1}^\top \mathbf{x} = \pm 2\sqrt{2} \tag{2.1.20}$$

Using (2.1.20) and (2.1.18),

$$\mathbf{x_m} = \begin{pmatrix} \pm 2\sqrt{2} \\ 4 \end{pmatrix} \tag{2.1.21}$$

Using these values of $\mathbf{x_m}$, the distances are

$$\left\| \begin{pmatrix} 0 \\ 5 \end{pmatrix} - \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right\| = 5 \tag{2.1.22}$$

$$\left\| \begin{pmatrix} 0 \\ 5 \end{pmatrix} - \begin{pmatrix} \pm 2\sqrt{2} \\ 4 \end{pmatrix} \right\| = 3 \tag{2.1.23}$$

Thus, the correct answer is **a)**.

2.2 Find the point on the curve

$$x^2 = 2y \tag{2.2.1}$$

which is nearest to the point $\mathbf{P} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$.

Solution:

(a) Quadratic Progamming

Using the relaxation in (2.2.3), the optimization problem can be framed as

$$\min_{\mathbf{x}} g(\mathbf{x}) = \|\mathbf{x} - \mathbf{P}\|^2 \tag{2.2.2}$$

$$\text{s.t. } h(\mathbf{x}) = \mathbf{x}^\top \mathbf{V}\mathbf{x} + 2\mathbf{u}^\top \mathbf{x} + f \leq 0 \tag{2.2.3}$$

where

$$\mathbf{V} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \ \mathbf{u} = \begin{pmatrix} 0 \\ -1 \end{pmatrix}, \ f = 0 \tag{2.2.4}$$

We can solve the above problem using quadratic programming (QP).

(b) Semi-definite Programming

We apply the semidefinite relaxation to get the following problem.

$$\min_{\mathbf{X}} \operatorname{tr}(\mathbf{C}\mathbf{X}) \tag{2.2.5}$$

$$\text{s.t. } \operatorname{tr}(\mathbf{A}\mathbf{X}) \leq 0 \tag{2.2.6}$$

$$\mathbf{X} \succeq 0 \tag{2.2.7}$$

where

$$\mathbf{C} = \begin{pmatrix} \mathbf{I} & -\mathbf{P} \\ -\mathbf{P}^\top & \|\mathbf{P}\|^2 \end{pmatrix} \tag{2.2.8}$$

$$\mathbf{A} = \begin{pmatrix} \mathbf{V} & \mathbf{u} \\ \mathbf{u}^\top & f \end{pmatrix} \tag{2.2.9}$$

The problem is solved using *cvxpy*.

(c) Lagrange Multipliers We need to find

$$\min_{\mathbf{x}} g(\mathbf{x}) = \|\mathbf{x} - \mathbf{P}\|^2 \tag{2.2.10}$$

$$\text{s.t. } h(\mathbf{x}) = \mathbf{x}^\top \mathbf{V}\mathbf{x} + 2\mathbf{u}^\top \mathbf{x} = 0 \tag{2.2.11}$$

where

$$\mathbf{V} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \ \mathbf{u} = \begin{pmatrix} 0 \\ -1 \end{pmatrix} \tag{2.2.12}$$

We use the method of Lagrange multipliers to find the optima. Here, we need to find $\lambda \in \mathbb{R}$ such that there exists a $\mathbf{x}$ satisfying

$$\nabla g(\mathbf{x}) = \lambda \nabla h(\mathbf{x}) \tag{2.2.13}$$

$$\implies 2(\mathbf{x} - \mathbf{P}) = 2\lambda(\mathbf{V}\mathbf{x} + \mathbf{u}) \tag{2.2.14}$$

$$\implies (\mathbf{I} - \lambda\mathbf{V})\mathbf{x} = \lambda\mathbf{u} + \mathbf{P} \tag{2.2.15}$$

$$\implies \begin{pmatrix} 1-\lambda & 0 \\ 0 & 1 \end{pmatrix} \mathbf{x} = \begin{pmatrix} 2 \\ 1-\lambda \end{pmatrix} \tag{2.2.16}$$

From (2.2.16), we have two cases:

i. $\lambda \neq 1$. In this case, we form the augmented matrix

$$\begin{pmatrix} 1-\lambda & 0 & 2 \\ 0 & 1 & 1-\lambda \end{pmatrix} \xleftrightarrow{R_1 \leftarrow \frac{R_1}{1-\lambda}} \begin{pmatrix} 1 & 0 & \frac{2}{1-\lambda} \\ 0 & 1 & 1-\lambda \end{pmatrix} \tag{2.2.17}$$

and get that

$$\mathbf{x_m} = \begin{pmatrix} \frac{2}{1-\lambda} \\ 1-\lambda \end{pmatrix} \tag{2.2.18}$$

Substituting in (2.2.11) with equality gives $\lambda = 1 - 2^{\frac{1}{3}}$. Thus,

$$\mathbf{x_m} = \begin{pmatrix} 2^{\frac{2}{3}} \\ 2^{\frac{1}{3}} \end{pmatrix} \tag{2.2.19}$$

ii. $\lambda = 1$. In this case, (2.2.16) becomes

$$\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \mathbf{x} = \begin{pmatrix} 2 \\ 0 \end{pmatrix} \tag{2.2.20}$$

which clearly has no solution.

Thus, the required point is

$$\mathbf{x_m} = \begin{pmatrix} 2^{\frac{2}{3}} \\ 2^{\frac{1}{3}} \end{pmatrix} \tag{2.2.21}$$

2.3 Find the equation of the normal to the curve $x^2 = 4y$ which passes through the point (4,-2)

Solution:

(a) Quadratic Progamming The given equation of the curve can be written as

$$g(\mathbf{x}) = \mathbf{x}^\top \mathbf{V}\mathbf{x} + 2\mathbf{u}^\top \mathbf{x} + f = 0 \tag{2.3.1}$$

where

$$\mathbf{V} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \mathbf{u} = \begin{pmatrix} 0 \\ -2 \end{pmatrix}, f = 0 \tag{2.3.2}$$

We are given that

$$\mathbf{P} = \begin{pmatrix} 4 \\ -2 \end{pmatrix} \tag{2.3.3}$$

This can be formulated as optimization problem as follows:

$$\min_{\mathbf{x}} \quad \mathrm{g}\left(\mathbf{x}\right) = \left\|\mathbf{x} - \mathbf{P}\right\|^2 \tag{2.3.4}$$

$$\text{s.t.} \quad h\left(\mathbf{x}\right) = \mathbf{x}^\top \mathbf{V}\mathbf{x} + 2\mathbf{u}^\top \mathbf{x} + f = 0 \tag{2.3.5}$$

From A.13, the given problem is not convex. Hence, using the relaxation (see A.14)

$$g\left(\mathbf{x}\right) = \mathbf{x}^\top \mathbf{V}\mathbf{x} + 2\mathbf{u}^\top \mathbf{x} + f \leq 0 \tag{2.3.6}$$

the optimization problem can be made convex. Solving the problem using *cvxpy* we get

$$\mathbf{x} = \begin{pmatrix} 1.695 \\ 0.718 \end{pmatrix} \tag{2.3.7}$$

See Fig. 2.3.1.

(b) Gradient Descent

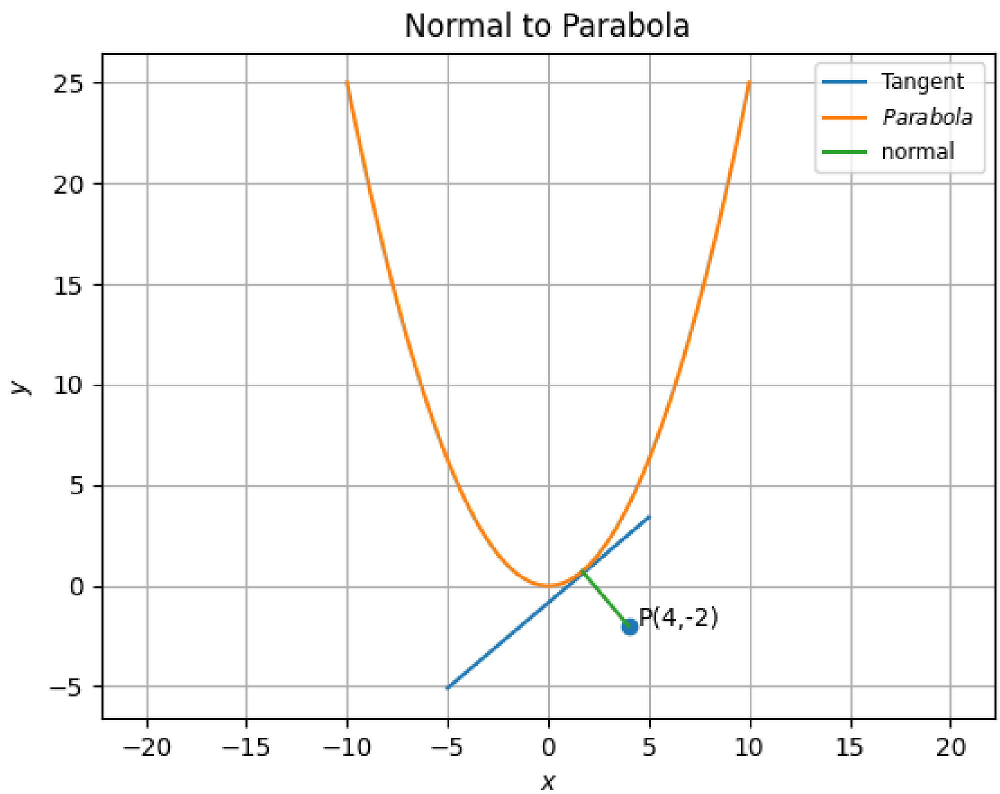


Figure 2.3.1:

Now we use Gradient Descent to find the optimum value. We define

$$\mathbf{x}_{n+1} = \mathbf{x}_n - \alpha \nabla g\left(\mathbf{x}_n\right) \tag{2.3.8}$$

And the condition is given as

$$\left(\mathbf{x} - \mathbf{h}\right)^\top \nabla g\left(\mathbf{x}_n\right) \neq 0 \tag{2.3.9}$$

Now we choose the parameters as

i. $\alpha = 0.001$

ii. precision $= 0.001$

iii. n $= 10000$

iv. $\mathbf{x}_0 = 4$

We get the minimum value of $\mathbf{x}$ as

$$\mathbf{x}_{min} = \begin{pmatrix} 1.695 \\ 0.718 \end{pmatrix} \tag{2.3.10}$$

(c) Semi-definite Programming The given equation of the curve can be written as

$$g\left(\mathbf{x}\right) = \mathbf{x}^\top \mathbf{V}\mathbf{x} + 2\mathbf{u}^\top \mathbf{x} + f = 0 \tag{2.3.11}$$

where

$$\mathbf{V} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \mathbf{u} = \begin{pmatrix} 0 \\ -2 \end{pmatrix}, f = 0 \tag{2.3.12}$$

We are given that

$$\mathbf{h} = \begin{pmatrix} 4 \\ -2 \end{pmatrix} \tag{2.3.13}$$

This can be formulated as optimization problem as below:

$$\min_{\mathbf{x}} \quad \mathrm{f}\left(\mathbf{x}\right) = \left\|\mathbf{x} - \mathbf{h}\right\|^2 \tag{2.3.14}$$

$$\text{s.t.} \quad g\left(\mathbf{x}\right) = \mathbf{x}^\top \mathbf{V}\mathbf{x} + 2\mathbf{u}^\top \mathbf{x} + f = 0 \tag{2.3.15}$$

Now,

$$\|\mathbf{x}-\mathbf{h}\|^2 = \|\mathbf{x}\|^2 - 2\mathbf{h}^\top\mathbf{x} + \|\mathbf{h}\|^2 \tag{2.3.16}$$

$$= \mathbf{y}^\top\mathbf{C}\mathbf{y} \tag{2.3.17}$$

where,

$$\mathbf{C} = \begin{pmatrix} \mathbf{I} & -\mathbf{h} \\ -\mathbf{h}^\top & \|h\|^2 \end{pmatrix} \text{ and } \mathbf{y} = \begin{pmatrix} \mathbf{x} \\ 1 \end{pmatrix} \tag{2.3.18}$$

And equation (2.3.15) can be expressed as

$$\mathbf{y}^\top\mathbf{A}\mathbf{y} = 0 \tag{2.3.19}$$

where

$$\mathbf{A} = \begin{pmatrix} \mathbf{V} & \mathbf{u} \\ \mathbf{u}^\top & f \end{pmatrix} \tag{2.3.20}$$

Using SDR(Semi Definite Relaxation), (2.3.14) can be expressed as

$$\min_{\mathbf{X}} tr\left(\mathbf{C}\mathbf{X}\right) \tag{2.3.21}$$

$$\text{s.t. } tr\left(\mathbf{A}\mathbf{X}\right) = 0 \tag{2.3.22}$$

$$\mathbf{X} \succcurlyeq \mathbf{0} \tag{2.3.23}$$

On solving it yields to the point

$$\mathbf{x} = \begin{pmatrix} 1.695 \\ 0.718 \end{pmatrix} \tag{2.3.24}$$

This is same as we obtained in Gradient Descent and Lagrange multiplier. Hence this is the required point.

(d) Lagrange Multipliers We will use Lagrange multipliers method to find the optimum value. Define

$$H(\mathbf{x}, \lambda) = f(\mathbf{x}) - \lambda g(\mathbf{x}) \tag{2.3.25}$$

and we find that

$$\nabla f(\mathbf{x}) = 2(\mathbf{x} - \mathbf{h}) \tag{2.3.26}$$

$$\nabla g(\mathbf{x}) = 2(\mathbf{V}\mathbf{x} + \mathbf{u}) \tag{2.3.27}$$

We have to find $\lambda \in \mathbb{R}$ such that

$$\nabla H\left(\mathbf{x}, \lambda\right) = 0 \tag{2.3.28}$$

$$\implies 2\left(\mathbf{x} - \mathbf{h}\right) - 2\lambda\left(\mathbf{V}\mathbf{x} + \mathbf{u}\right) = 0 \tag{2.3.29}$$

$$\implies \mathbf{x} - \mathbf{h} = \lambda\left(\mathbf{V}\mathbf{x} + \mathbf{u}\right) \tag{2.3.30}$$

$$\implies \left(\mathbf{I} - \lambda\mathbf{V}\right)\mathbf{x} = \lambda\mathbf{u} + \mathbf{h} \tag{2.3.31}$$

$$\implies \begin{pmatrix} 1-\lambda & 0 \\ 0 & 1 \end{pmatrix}\mathbf{x} = \lambda\begin{pmatrix} 0 \\ -2 \end{pmatrix} + \begin{pmatrix} 4 \\ -2 \end{pmatrix} \tag{2.3.32}$$

$$\implies \begin{pmatrix} 1-\lambda & 0 \\ 0 & 1 \end{pmatrix}\mathbf{x} = \begin{pmatrix} 4 \\ -2\lambda - 2 \end{pmatrix} \tag{2.3.33}$$

Writing augmented matrix,

$$\begin{pmatrix} 1-\lambda & 0 & 4 \\ 0 & 1 & -2\lambda-2 \end{pmatrix} \xleftrightarrow{R_1 \leftarrow \frac{R_1}{1-\lambda}} \begin{pmatrix} 1 & 0 & \frac{4}{1-\lambda} \\ 0 & 1 & -2\lambda-2 \end{pmatrix} \tag{2.3.34}$$

Then, we get

$$\mathbf{x}_m = \begin{pmatrix} \frac{4}{1-\lambda} \\ -2\lambda - 2 \end{pmatrix} \tag{2.3.35}$$

Substituting this value in (2.3.1)

$$\begin{pmatrix} \frac{4}{1-\lambda} & -2-2\lambda \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \frac{4}{1-\lambda} \\ -2-2\lambda \end{pmatrix} + 2 \begin{pmatrix} 0 & -2 \end{pmatrix} \begin{pmatrix} \frac{4}{1-\lambda} \\ -2-2\lambda \end{pmatrix} = 0 \quad (2.3.36)$$

$$\frac{16}{(1-\lambda)^2} + 8(\lambda+1) = 0 \quad (2.3.37)$$

$$\lambda^3 - \lambda^2 - \lambda + 3 = 0 \quad (2.3.38)$$

$$\implies \lambda = -1.3593 \quad (2.3.39)$$

Substituting the value of λ in (2.3.35)

$$\mathbf{x}_m = \begin{pmatrix} 1.695 \\ 0.718 \end{pmatrix} \quad (2.3.40)$$

This result is same as we obtained using gradient descent. Hence, it is the point of normal.

2.4 Find the normal to the curve $2y + x^2 = 3$ passing through (2,2).

Solution: The parameters of the given conic are

$$\mathbf{V} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \mathbf{u} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, f = -3 \quad (2.4.1)$$

If $\mathbf{x}$ be the point of contact on the conic, the optimization problem can be formulated as

$$\mathbf{q} = \min_{\mathbf{x}} \|\mathbf{x} - \mathbf{p}\|^2 \tag{2.4.2}$$

$$s.t. \quad \mathbf{x}^\top \mathbf{V}\mathbf{x} + 2\mathbf{u}^\top \mathbf{x} + f = 0 \tag{2.4.3}$$

where

$$\mathbf{p} = \begin{pmatrix} 2 \\ 2 \end{pmatrix} \tag{2.4.4}$$

Since

$$\|\mathbf{x} - \mathbf{p}\|^2 = \|\mathbf{x}\|^2 - 2\mathbf{p}^\top \mathbf{x} + \|\mathbf{p}\|^2 \tag{2.4.5}$$

$$= \mathbf{y}^\top \mathbf{C}\mathbf{y} \tag{2.4.6}$$

where

$$\mathbf{C} = \begin{pmatrix} \mathbf{I} & -\mathbf{p} \\ -\mathbf{p}^\top & \|\mathbf{p}\|^2 \end{pmatrix} \mathbf{y} = \begin{pmatrix} \mathbf{x} \\ 1 \end{pmatrix} \tag{2.4.7}$$

and (2.4.3) can be expressed as

$$\mathbf{y}^\top \mathbf{A}\mathbf{y} = 0, \tag{2.4.8}$$

where

$$\mathbf{A} = \begin{pmatrix} \mathbf{V} & \mathbf{u} \\ \mathbf{u}^\top & f \end{pmatrix}, \tag{2.4.9}$$

Using SDR (Semi Definite Relaxation), (2.4.2) can be expressed as

$$\min_{\mathbf{X}} tr\left(\mathbf{CX}\right) \tag{2.4.10}$$

$$s.t. \quad tr\left(\mathbf{AX}\right) = 0, \tag{2.4.11}$$

$$\mathbf{X} \succeq \mathbf{0} \tag{2.4.12}$$

yielding

$$\mathbf{x} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}. \tag{2.4.13}$$

Thus, the equation of the normal is

$$\begin{pmatrix} 1 & -1 \end{pmatrix} \mathbf{x} = 0 \tag{2.4.14}$$

2.5 Find the equation of the normal to the curve $x^2 = 4y$ and passing through the point $(1, 2)$.

Solution:

(a) Optimization Problem

The given equation of the curve can be written as

$$g\left(\mathbf{x}\right) = \mathbf{x}^T\mathbf{V}\mathbf{x} + 2\mathbf{u}^T\mathbf{x} + f = 0 \tag{2.5.1}$$

The given problem can be formulated as

$$\min_{\mathbf{x}} \quad \mathrm{f}(\mathbf{x}) = \|\mathbf{x} - \mathbf{h}\|^2 \tag{2.5.2}$$

$$\text{s.t.} \quad g(\mathbf{x}) = \mathbf{x}^T\mathbf{V}\mathbf{x} + 2\mathbf{u}^T\mathbf{x} + f = 0 \tag{2.5.3}$$

where

$$\mathbf{V} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \mathbf{u} = \begin{pmatrix} 0 \\ -2 \end{pmatrix}, f = 0, \mathbf{h} = \begin{pmatrix} 1 \\ 2 \end{pmatrix} \tag{2.5.4}$$

The optimization problem is nonconvex. However, by relaxing the constraint in (2.5.3) as

$$g(\mathbf{x}) = \mathbf{x}^T\mathbf{V}\mathbf{x} + 2\mathbf{u}^T\mathbf{x} + f \le 0 \tag{2.5.5}$$

the optimization problem can be made convex. Using cvxpy, input the objective function, contraints and solve. However, resultant optimal point is the given point itself. This is because, the point is inside the parabola. Looks like, this is a limitation of cvxpy.

(b) Semi-definite Programming Find the normal to the curve $x^2 = 4y$ passing through (1,2). The parameters of the given conic are

$$\mathbf{V} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \mathbf{u} = \begin{pmatrix} 0 \\ -2 \end{pmatrix}, f = 0 \tag{2.5.6}$$

If $\mathbf{x}$ be the point of contact on the conic, the optimization problem can be formulated as

$$\mathbf{q} = \min_{\mathbf{x}} \|\mathbf{x} - \mathbf{p}\|^2 \tag{2.5.7}$$

$$s.t. \quad \mathbf{x}^\top \mathbf{V} \mathbf{x} + 2\mathbf{u}^\top \mathbf{x} + f = 0 \tag{2.5.8}$$

where

$$\mathbf{p} = \begin{pmatrix} 1 \\ 2 \end{pmatrix} \tag{2.5.9}$$

Since

$$\|\mathbf{x} - \mathbf{p}\|^2 = \|\mathbf{x}\|^2 - 2\mathbf{p}^\top \mathbf{x} + \|\mathbf{p}\|^2 \tag{2.5.10}$$

$$= \mathbf{y}^\top \mathbf{C} \mathbf{y} \tag{2.5.11}$$

where

$$\mathbf{C} = \begin{pmatrix} \mathbf{I} & -\mathbf{p} \\ -\mathbf{p}^\top & \|\mathbf{p}\|^2 \end{pmatrix} \mathbf{y} = \begin{pmatrix} \mathbf{x} \\ 1 \end{pmatrix} \tag{2.5.12}$$

and (2.5.8) can be expressed as

$$\mathbf{y}^\top \mathbf{A} \mathbf{y} = 0, \tag{2.5.13}$$

where

$$\mathbf{A} = \begin{pmatrix} \mathbf{V} & \mathbf{u} \\ \mathbf{u}^\top & f \end{pmatrix}, \tag{2.5.14}$$

Using SDR (Semi Definite Relaxation), (2.5.7) can be expressed as

$$\min_{\mathbf{X}} tr\left(\mathbf{C}\mathbf{X}\right) \tag{2.5.15}$$

$$s.t. \quad tr\left(\mathbf{A}\mathbf{X}\right) = 0, \tag{2.5.16}$$

$$\mathbf{X} \succeq \mathbf{0} \tag{2.5.17}$$

yielding

$$\mathbf{x} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}. \tag{2.5.18}$$

Thus, the equation of the normal is

$$\begin{pmatrix} 1 & 1 \end{pmatrix} \mathbf{x} = 3 \tag{2.5.19}$$

(c) Lagrange Multipliers

The given equation of the curve can be written as

$$g\left(\mathbf{x}\right) = \mathbf{x}^T \mathbf{V} \mathbf{x} + 2\mathbf{u}^T \mathbf{x} + f = 0 \tag{2.5.20}$$

where

$$\mathbf{V} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \tag{2.5.21}$$

$$\mathbf{u} = \begin{pmatrix} 0 \\ -2 \end{pmatrix} \tag{2.5.22}$$

$$f = 0 \tag{2.5.23}$$

We are given that

$$\mathbf{h} = \begin{pmatrix} 1 \\ 2 \end{pmatrix} \tag{2.5.24}$$

This can be formulated as optimization problem as below:

$$\min_{\mathbf{x}} \quad \mathrm{f}\left(\mathbf{x}\right) = \left\|\mathbf{x} - \mathbf{h}\right\|^2 \tag{2.5.25}$$

$$\text{s.t.} \quad g\left(\mathbf{x}\right) = \mathbf{x}^T\mathbf{V}\mathbf{x} + 2\mathbf{u}^T\mathbf{x} + f = 0 \tag{2.5.26}$$

It is already proved that the optimization problem is nonconvex. The constraints throw an error when *cvxpy* is used. We will use Lagrange multipliers method to find the optimum value. Define

$$H\left(\mathbf{x}, \lambda\right) = f\left(\mathbf{x}\right) - \lambda g\left(\mathbf{x}\right) \tag{2.5.27}$$

and we find that

$$\nabla f(\mathbf{x}) = 2(\mathbf{x} - \mathbf{h}) \tag{2.5.28}$$

$$\nabla g(\mathbf{x}) = 2(\mathbf{V}\mathbf{x} + \mathbf{u}) \tag{2.5.29}$$

We have to find $\lambda \in \mathbb{R}$ such that

$$\nabla H(\mathbf{x}, \lambda) = 0 \tag{2.5.30}$$

$$\implies 2(\mathbf{x} - \mathbf{h}) - 2\lambda(\mathbf{V}\mathbf{x} + \mathbf{u}) = 0 \tag{2.5.31}$$

$$\implies \mathbf{x} - \mathbf{h} = \lambda(\mathbf{V}\mathbf{x} + \mathbf{u}) \tag{2.5.32}$$

$$\implies (\mathbf{I} - \lambda\mathbf{V})\mathbf{x} = \lambda\mathbf{u} + \mathbf{h} \tag{2.5.33}$$

$$\implies \begin{pmatrix} 1-\lambda & 0 \\ 0 & 1 \end{pmatrix}\mathbf{x} = \lambda\begin{pmatrix} 0 \\ -2 \end{pmatrix} + \begin{pmatrix} 1 \\ 2 \end{pmatrix} \tag{2.5.34}$$

$$\implies \begin{pmatrix} 1-\lambda & 0 \\ 0 & 1 \end{pmatrix}\mathbf{x} = \begin{pmatrix} 1 \\ -2\lambda + 2 \end{pmatrix} \tag{2.5.35}$$

We have 2 cases to considers here.

i. When $\lambda \neq 1$. Writing augmented matrix,

$$\begin{pmatrix} 1-\lambda & 0 & 1 \\ 0 & 1 & -2\lambda + 2 \end{pmatrix} \xleftrightarrow{R_1 \leftarrow \frac{R_1}{1-\lambda}} \begin{pmatrix} 1 & 0 & \frac{1}{1-\lambda} \\ 0 & 1 & -2\lambda + 2 \end{pmatrix} \tag{2.5.36}$$

Then, we get

$$\mathbf{x}_m = \begin{pmatrix} \frac{1}{1-\lambda} \\ -2\lambda + 2 \end{pmatrix} \tag{2.5.37}$$

Substituting this value in (2.5.26)

$$\begin{pmatrix} \frac{1}{1-\lambda} & -2\lambda+2 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \frac{1}{1-\lambda} \\ -2\lambda+2 \end{pmatrix} + 2 \begin{pmatrix} 0 & -2 \end{pmatrix} \begin{pmatrix} \frac{1}{1-\lambda} \end{pmatrix} = 0$$

$$\implies (\lambda - 1)^3 = -\frac{1}{8}$$

$$\implies \lambda = \frac{1}{2} \quad (2.5.38)$$

Substituting the value of λ in (2.5.37)

$$\mathbf{x}_m = \mathbf{q} = \begin{pmatrix} \frac{1}{1-\frac{1}{2}} \\ -2\left(\frac{1}{2}\right)+2 \end{pmatrix} \quad (2.5.39)$$

$$= \begin{pmatrix} 2 \\ 1 \end{pmatrix} \quad (2.5.40)$$

ii. When $\lambda = 1$, from (2.5.35),

$$\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \mathbf{x} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad (2.5.41)$$

This is an invalid solution.

Given the point of contact $\mathbf{q}$, the equation to the normal is given by

$$\left(\mathbf{V}\mathbf{q}+\mathbf{u}\right)^{\top}\mathbf{R}\left(\mathbf{x}-\mathbf{q}\right)=0 \tag{2.5.42}$$

$$\implies \left(\begin{pmatrix}1 & 0\\0 & 0\end{pmatrix}\begin{pmatrix}2\\1\end{pmatrix}+\begin{pmatrix}0\\-2\end{pmatrix}\right)^{\top}\begin{pmatrix}0 & 1\\-1 & 0\end{pmatrix}\left(\mathbf{x}-\begin{pmatrix}2\\1\end{pmatrix}\right)=0 \tag{2.5.43}$$

$$\implies \begin{pmatrix}1 & 1\end{pmatrix}\mathbf{x}=3 \tag{2.5.44}$$

The relevant figure is shown in 2.5.1

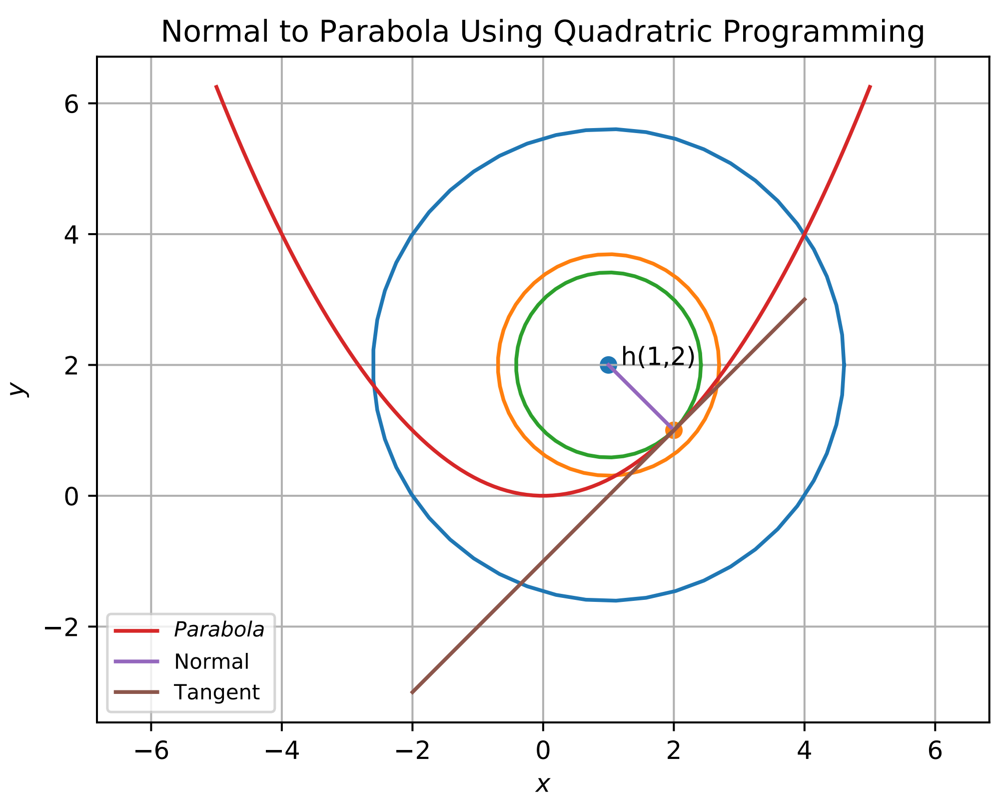


Figure 2.5.1:

Chapter 3

Geometric Programming

3.1. Examples

3.1 Show that the semi-vertical angle of the cone of the maximum volume and of given slant height is $\tan^{-1}\sqrt{2}$.

Solution: We use geometric programming. Taking the radius to be r, height to be h, and slant height $; = 1$ without loss of generality, we need to find

$$\max_{r,h} \frac{1}{3}\pi r^2 h \tag{3.1.1}$$

$$\text{s.t. } r^2 + h^2 = 1 \tag{3.1.2}$$

$$r, h \geq 0 \tag{3.1.3}$$

The Python code solves this Disciplined Geometric Programming (DGP) problem using *cvxpy*. The solutions are

$$r_M = \sqrt{\frac{2}{3}},\ h_M = \frac{1}{\sqrt{3}} \tag{3.1.4}$$

Hence, from (3.1.4), the required semi-verical angle is

$$\alpha = \tan^{-1}\frac{r}{h} = \tan^{-1}\sqrt{2} \tag{3.1.5}$$

as required.

3.2 Show that semi-vertical angle of right circular cone of given surface area and maximum volume is $\sin^{-1}\left(\frac{1}{3}\right)$.

Solution: Let r, h, l be the radius, height and slant height of the right circular cone respectively. Let S be the given surface area and V be the volume of the cone. We have

$$l^2 = r^2 + h^2 \tag{3.2.1}$$

$$S = \pi rl + \pi r^2 \tag{3.2.2}$$

$$\implies l = \frac{S - \pi r^2}{\pi r} \tag{3.2.3}$$

$$V = \frac{1}{3}\pi r^2 h \tag{3.2.4}$$

The given problem can be formulated as

$$V = \max_{r,h} \frac{1}{3}\pi r^2 h \tag{3.2.5}$$

$$\text{s.t } \pi r\left(\sqrt{r^2 + h^2}\right) + \pi r^2 \leq S \tag{3.2.6}$$

(a) Theoritical proof: From (3.2.4),

$$V = \frac{1}{3}\pi r^2\sqrt{l^2 - r^2} \tag{3.2.7}$$

$$\implies V^2 = \frac{1}{9}\pi^2 r^4\left(l^2 - r^2\right) \tag{3.2.8}$$

$$= \frac{1}{9}\left(S^2 r^2 - 2\pi S r^4\right) \tag{3.2.9}$$

Differentiating wrt r,

$$2V\frac{dV}{dr} = \frac{S^2}{9}2r - \frac{2\pi S}{9}4r^3 \tag{3.2.10}$$

$$= \frac{2rS}{9}\left(S - 4\pi r^2\right) \tag{3.2.11}$$

For maximum volume, $\frac{dV}{dr} = 0$

$$\implies \frac{2rS}{9}\left(S - 4\pi r^2\right) = 0 \tag{3.2.12}$$

$$\implies r = 0 \text{ or } S - 4\pi r^2 = 0 \tag{3.2.13}$$

Since $r \neq 0$,

$$S - 4\pi r^2 = 0 \tag{3.2.14}$$

$$\implies r^2 = \frac{S}{4\pi} = \frac{\pi r l + \pi r^2}{4\pi} \tag{3.2.15}$$

$$\implies l = 3r \tag{3.2.16}$$

For V to be maximum, $\frac{d^2V}{dr^2} < 0$, from (3.2.10)

$$\frac{dV}{dr} = \frac{S}{3}\left[\frac{S - 4\pi r^2}{\sqrt{S^2 - 2\pi S r^2}}\right] \tag{3.2.17}$$

$$\implies \frac{d^2V}{dr^2} = \frac{S}{3}\left[\frac{8\pi^2 S r^3 - 6\pi r S^2}{(S^2 - 2\pi S r^2)^{\frac{3}{2}}}\right] \tag{3.2.18}$$

For maximum volume, substituting the value of S from (3.2.15)

$$\frac{d^2V}{dr^2} = \frac{d^2V}{dr^2} < 0 \tag{3.2.19}$$

Let θ be the semi-vertical angle in Figure 3.2.1. Then,

$$\sin\theta = \frac{OA}{CA} = \frac{r}{l} \tag{3.2.20}$$

$$\sin\theta = \frac{r}{3r} \tag{3.2.21}$$

$$\implies \theta = \sin^{-1}\frac{1}{3} \tag{3.2.22}$$

(b) Using Disciplined Geometric Programming (DGP) of cvxpy: Refer to equations (3.2.5) and (3.2.6) for formulation of optimization problem. Assume S to be 75.42857 sq.units. Solving this problem, yields following results:

$$r = 2.45 \text{ units} \tag{3.2.23}$$

$$l = 7.35 \text{ units} \tag{3.2.24}$$

$$h = 6.93 \text{ units} \tag{3.2.25}$$

$$\text{Optimal } V \approx 43.557 \text{ cu.units} \tag{3.2.26}$$

It can be seen from solution that $l = 3r$ and semi-vertical angle is given as

$\sin^{-1}\left(\frac{1}{3}\right)$. This is similar to what we proved theoritically.

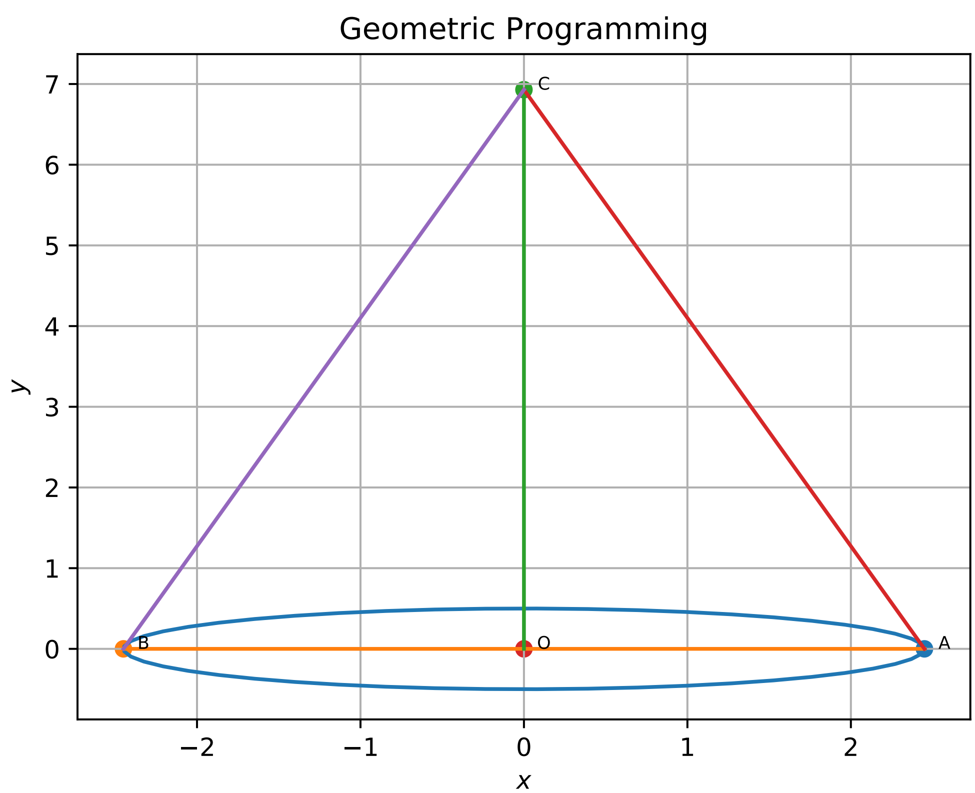


Figure 3.2.1:

3.3 Find the maximum area of an isosceles triangle inscribed in the ellipse $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$ with its vertex at one end of the major axis.

Solution:

Chapter 4

Linear Programming

4.1. Definition

4.1.1 Maximize

$$Z = 3x + 4y \tag{4.1.1.1}$$

subject to the constraints:

$$x + 4y \leq 4, \tag{4.1.1.2}$$

$$x \geq 0, y \geq 0 \tag{4.1.1.3}$$

Solution:

(a) Using cvxpy method: The given problem can be formulated as

$$\max_{\mathbf{x}} Z = \begin{pmatrix} 3 & 4 \end{pmatrix} \mathbf{x} \tag{4.1.1.4}$$

$$\begin{pmatrix} 1 & 4 \\ -1 & 0 \\ 0 & -1 \end{pmatrix} \mathbf{x} \preceq \begin{pmatrix} 4 \\ 0 \\ 0 \end{pmatrix} \tag{4.1.1.5}$$

Solving using cvxpy, we get

$$\max_{\mathbf{x}} Z = 12, \mathbf{x} = \begin{pmatrix} 4 \\ 0 \end{pmatrix} \tag{4.1.1.6}$$

(b) Using Corner point method: The corner points of the inequalities are:

$$\mathbf{A} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{4.1.1.7}$$

$$\mathbf{B} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \tag{4.1.1.8}$$

$$\mathbf{x} = \begin{pmatrix} 4 \\ 0 \end{pmatrix} \tag{4.1.1.9}$$

Substituting above values of corner points in Equation (4.1.1.1) to get the value of Z, as shown in the Table 4.1.1.2

Corner Point	**Corresponding Z value**
$\mathbf{A} \begin{pmatrix} 0 \\ 1 \end{pmatrix}$	4
$\mathbf{B} \begin{pmatrix} 0 \\ 0 \end{pmatrix}$	0
$\mathbf{x} \begin{pmatrix} 4 \\ 0 \end{pmatrix}$	12

Table 4.1.1.2:

From the table 4.1.1.2, it is clear that the optimum value and optimum point are similar to what we found in (4.1.1.6).

The relevant figure is as shown in 4.1.1.1

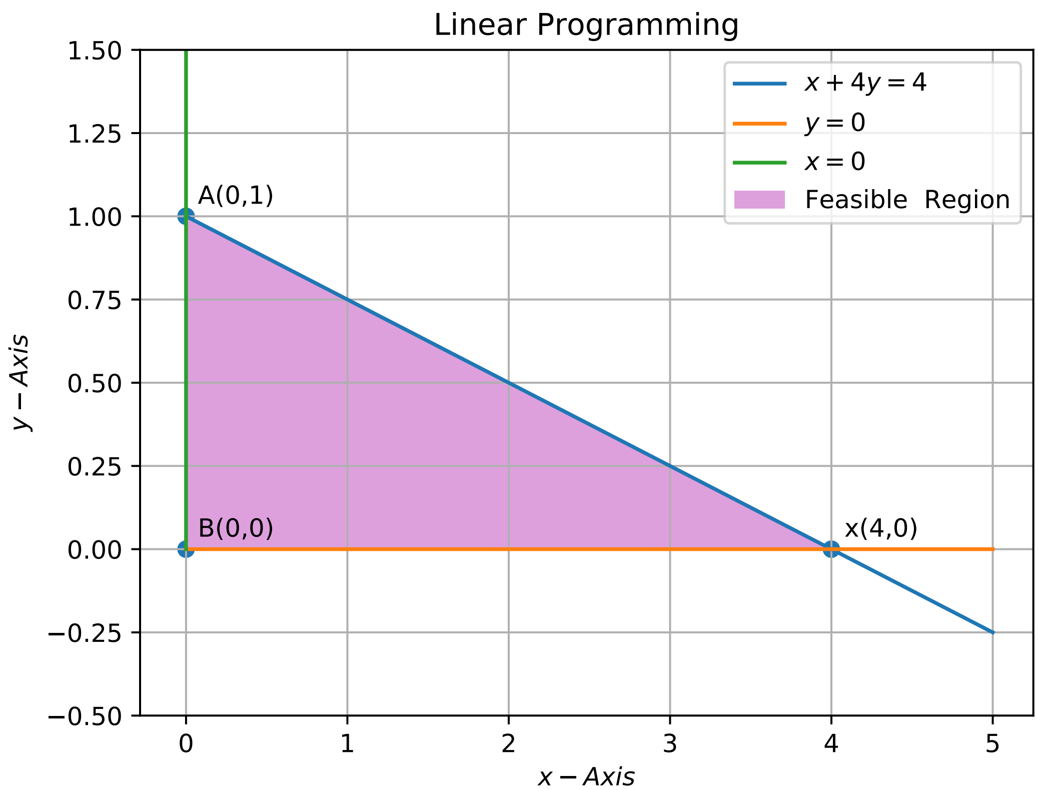


Figure 4.1.1.1:

4.1.2 Minimise

$$Z = -3x + 4y, \tag{4.1.2.1}$$

subject to $x + 2y \leq 8, 3x + 2y \leq 12, x \geq 0, y \geq 0$.

Solution:

(a) Using cvxpy method: The given problem can be formulated as

$$\min_{\mathbf{x}} Z = \begin{pmatrix} -3 & 4 \end{pmatrix} \mathbf{x} \tag{4.1.2.2}$$

$$\text{s.t.} \quad \mathbf{Ax} \leq B \tag{4.1.2.3}$$

where

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 2 \\ -1 & 0 \\ 0 & -1 \end{pmatrix}, B = \begin{pmatrix} 8 \\ 12 \\ 0 \\ 0 \end{pmatrix} \tag{4.1.2.4}$$

By solving using cvxpy, we get

$$\min_{\mathbf{x}} Z = -12 \tag{4.1.2.5}$$

$$\mathbf{x} = \begin{pmatrix} 4 \\ 0 \end{pmatrix} \tag{4.1.2.6}$$

(b) Using Corner point method: See Fig. 4.1.2.1. The corner points of the inequalities are:

$$\mathbf{P} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \mathbf{Q} = \begin{pmatrix} 0 \\ 4 \end{pmatrix}, \mathbf{R} = \begin{pmatrix} 2 \\ 3 \end{pmatrix}, \mathbf{S} = \begin{pmatrix} 4 \\ 0 \end{pmatrix} \tag{4.1.2.7}$$

We have

$$Z = -3x + 4y \tag{4.1.2.8}$$

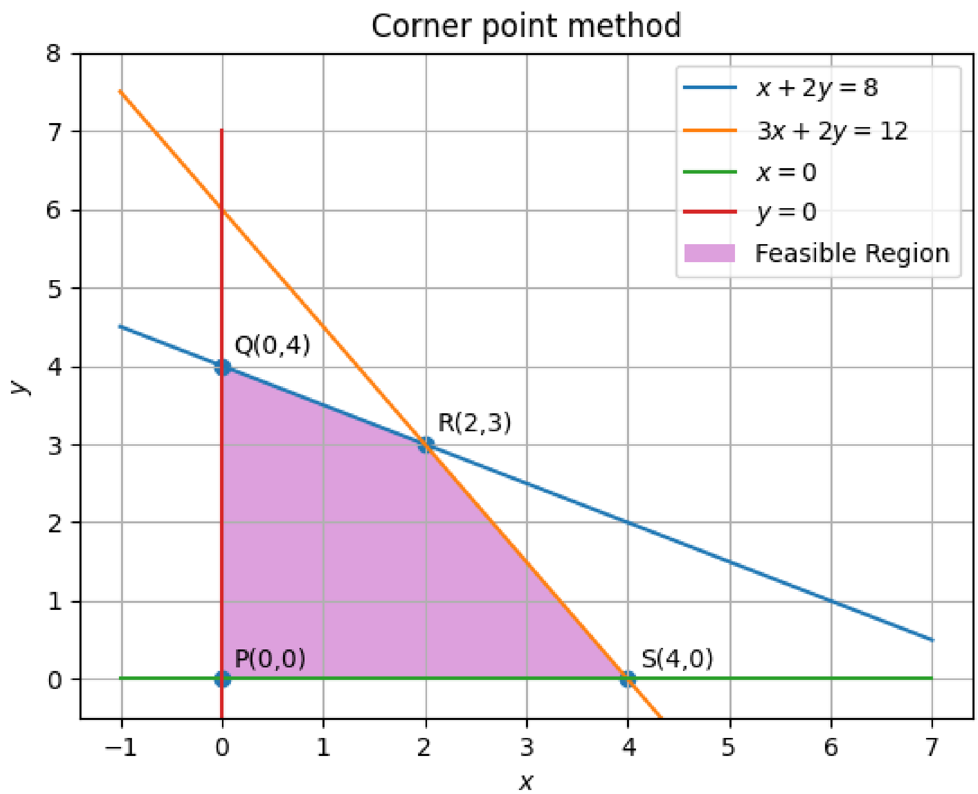


Figure 4.1.2.1: Graph

Substituting above values of corner points in (4.1.2.8) to get the value of Z, as shown in Table 4.1.2.1, the optimum point and optimum value are

$$\mathbf{S} = \begin{pmatrix} 4 \\ 0 \end{pmatrix} \tag{4.1.2.9}$$

$$\min Z = -12 \tag{4.1.2.10}$$

Corner Point	**Corresponding Z value**
P $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$	0
Q $\begin{pmatrix} 0 \\ 4 \end{pmatrix}$	16
R $\begin{pmatrix} 2 \\ 3 \end{pmatrix}$	6
S $\begin{pmatrix} 4 \\ 0 \end{pmatrix}$	-12

Table 4.1.2.1:

4.1.3 Maximize $Z = 5x + 3y$ such that

$$3x + 5y \leq 15, \tag{4.1.3.1}$$

$$5x + 2y \leq 10, \tag{4.1.3.2}$$

$$x \geq 0, y \geq 0. \tag{4.1.3.3}$$

Solution: The given problem can be expressed as

$$Z = \begin{pmatrix} 5 & 3 \end{pmatrix} \mathbf{x} \tag{4.1.3.4}$$

$$s.t. \quad \begin{pmatrix} 3 & 5 \\ 5 & 2 \end{pmatrix} \mathbf{x} \leq \begin{pmatrix} 15 \\ 10 \end{pmatrix} \tag{4.1.3.5}$$

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \mathbf{x} \geq \begin{pmatrix} 0 \\ 0 \end{pmatrix} \tag{4.1.3.6}$$

$$\tag{4.1.3.7}$$

Using cvxpy, the solution is

$$\mathbf{x} = \begin{pmatrix} \frac{20}{19} \\ \frac{45}{19} \end{pmatrix}, Z_{max} = \frac{235}{19} \tag{4.1.3.8}$$

4.1.4 Minimize $Z = 3x + 5y$ such that

$$x + 3y \geq 3 \tag{4.1.4.1}$$

$$x + y \geq 2 \tag{4.1.4.2}$$

$$x \geq 0, y \geq 0. \tag{4.1.4.3}$$

Solution: The given problem can be expressed as

$$Z = \min_{\mathbf{x}} \begin{pmatrix} 3 & 5 \end{pmatrix} \mathbf{x} \tag{4.1.4.4}$$

$$\begin{pmatrix} 1 & 3 \\ 1 & 1 \end{pmatrix} \mathbf{x} \succeq \begin{pmatrix} 3 \\ 2 \end{pmatrix} \tag{4.1.4.5}$$

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \mathbf{x} \succeq \begin{pmatrix} 0 \\ 0 \end{pmatrix} \tag{4.1.4.6}$$

Solving using cvxpy, we get,

$$\mathbf{x} = \begin{pmatrix} \frac{3}{2} \\ \frac{1}{2} \end{pmatrix}, Z_{min} = 7 \tag{4.1.4.7}$$

4.1.5

4.1.6 Minimize Z=x+2y subject to

$$2x + 3y \geq 3 \tag{4.1.6.1}$$

$$x + 2y \geq 6 \tag{4.1.6.2}$$

$$x, y \geq 0. \tag{4.1.6.3}$$

Solution: The optimization problem can be defined as

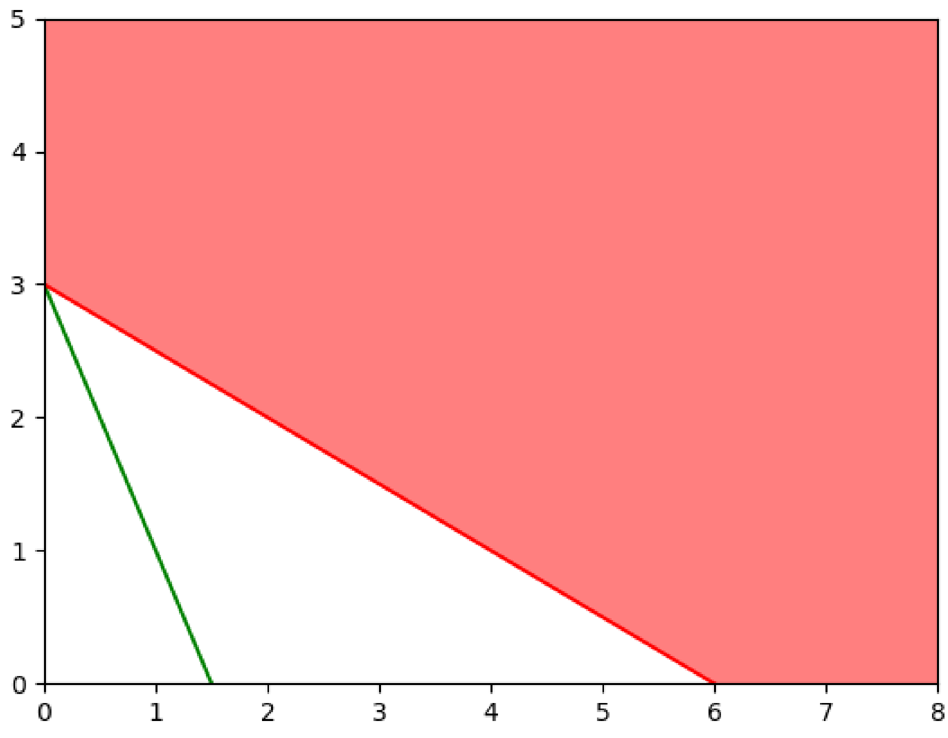


Figure 4.1.6.1:

$$P = \min_{\mathbf{x}} \begin{pmatrix} 1 & 2 \end{pmatrix} \mathbf{x} \tag{4.1.6.4}$$

$$\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} \mathbf{x} \succeq \begin{pmatrix} 3 \\ 6 \end{pmatrix} \tag{4.1.6.5}$$

$$x, y \geq \mathbf{0} \tag{4.1.6.6}$$

From Fig. 4.1.6.1, the feasible region vertices are

$$\begin{pmatrix} 0 \\ 3 \end{pmatrix}, \begin{pmatrix} 6 \\ 0 \end{pmatrix} \tag{4.1.6.7}$$

yielding

$$\begin{pmatrix} 1 & 2 \end{pmatrix} \begin{pmatrix} 0 \\ 3 \end{pmatrix} = 6 \tag{4.1.6.8}$$

$$\begin{pmatrix} 1 & 2 \end{pmatrix} \begin{pmatrix} 6 \\ 0 \end{pmatrix} = 6 \tag{4.1.6.9}$$

$$\tag{4.1.6.10}$$

Thus, the minimum value of Z is 6.

4.1.7 Minimise and Maximise

$$Z = 5x + 10y \tag{4.1.7.1}$$

subject to

$$x + 2y \leq 120 \tag{4.1.7.2}$$

$$x + y \geq 60 \tag{4.1.7.3}$$

$$x - 2y \geq 0 \tag{4.1.7.4}$$

$$x \geq 0, y \geq 0 \tag{4.1.7.5}$$

Solution: The given problem can be formulated as

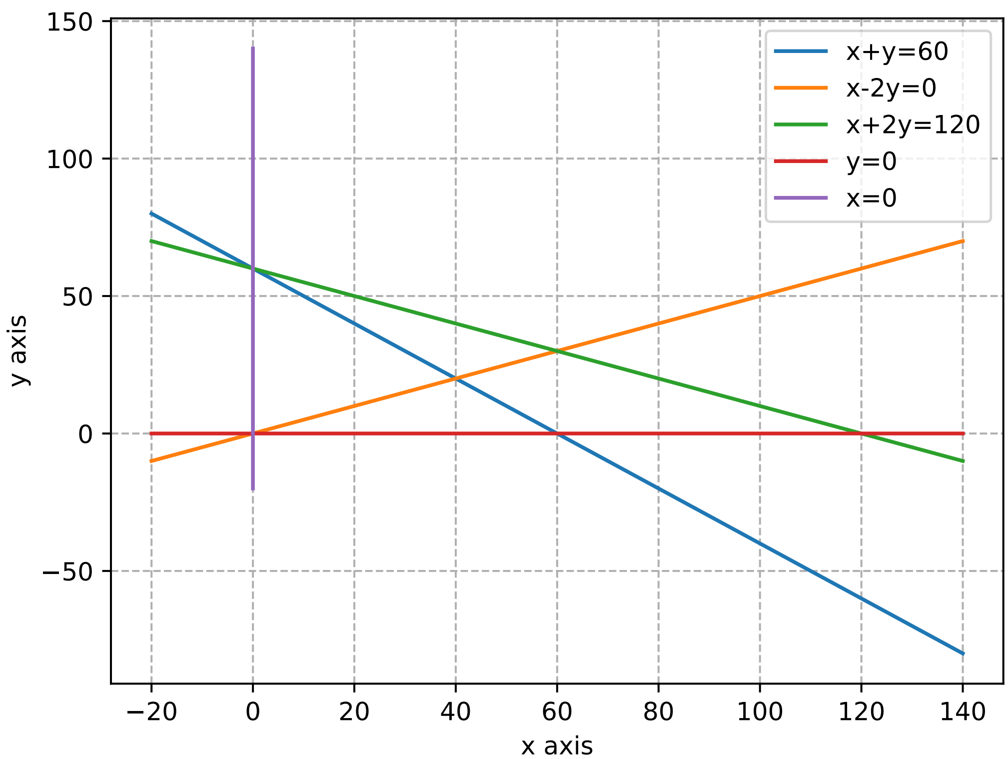


Figure 4.1.7.1:

$$\min_{\mathbf{x}} \mathbf{Z} = \begin{pmatrix} 5 & 10 \end{pmatrix} \mathbf{x} \tag{4.1.7.6}$$

$$\max_{\mathbf{x}} \mathbf{Z} = \begin{pmatrix} 5 & 10 \end{pmatrix} \mathbf{x} \tag{4.1.7.7}$$

$$s.t. \quad \begin{pmatrix} -1 & -2 \\ 1 & 1 \\ 1 & -2 \\ 1 & 0 \\ 0 & 1 \end{pmatrix} \mathbf{x} \succeq \begin{pmatrix} -120 \\ 60 \\ 0 \\ 0 \\ 0 \end{pmatrix} \tag{4.1.7.8}$$

$$\tag{4.1.7.9}$$

Solving above equations using cvxpy,

$$\min_{\mathbf{x}} Z = 300, \mathbf{x} = \begin{pmatrix} 60 \\ 0 \end{pmatrix} \tag{4.1.7.10}$$

$$\max_{\mathbf{x}} Z = 600, \mathbf{x} = \begin{pmatrix} 60 \\ 30 \end{pmatrix} \tag{4.1.7.11}$$

4.1.8

4.1.9 Maximise

$$Z = -x + 2y \tag{4.1.9.1}$$

subject to the constraints

$$x + y \geq 5 \tag{4.1.9.2}$$

$$x + 2y \geq 6 \tag{4.1.9.3}$$

$$x \geq 3, y \geq 0. \tag{4.1.9.4}$$

Solution: The given problem can be expressed as

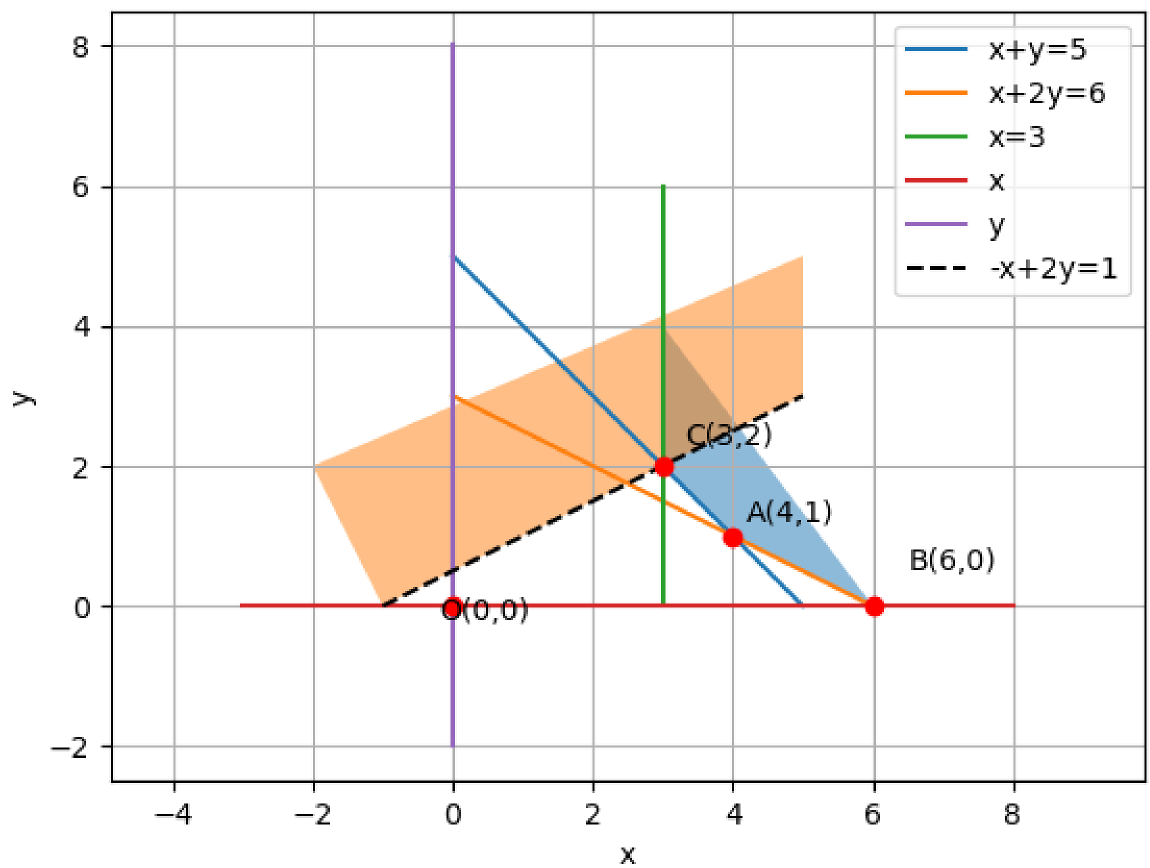


Figure 4.1.9.1:

$$z = \max_{\mathbf{x}} \begin{pmatrix} -1 & 2 \end{pmatrix} \mathbf{x} \tag{4.1.9.5}$$

$$s.t. \quad \begin{pmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 0 \\ 0 & 1 \end{pmatrix} \mathbf{x} = \begin{pmatrix} 5 \\ 6 \\ 3 \\ 0 \end{pmatrix} \tag{4.1.9.6}$$

By providing the objective function and constraints to cvxpy, the optimal value gives infinity as result and the problem is unbounded. This is verified from Fig. 4.1.9.1.

4.1.10 Maximize

$$Z = x + y \tag{4.1.10.1}$$

subject to

$$x - y \leq -1 \tag{4.1.10.2}$$

$$-x + y \leq 0 \tag{4.1.10.3}$$

$$x, y \geq 0 \tag{4.1.10.4}$$

Solution: From Fig. 4.1.10.1, the given problem has no optimal solution. This is

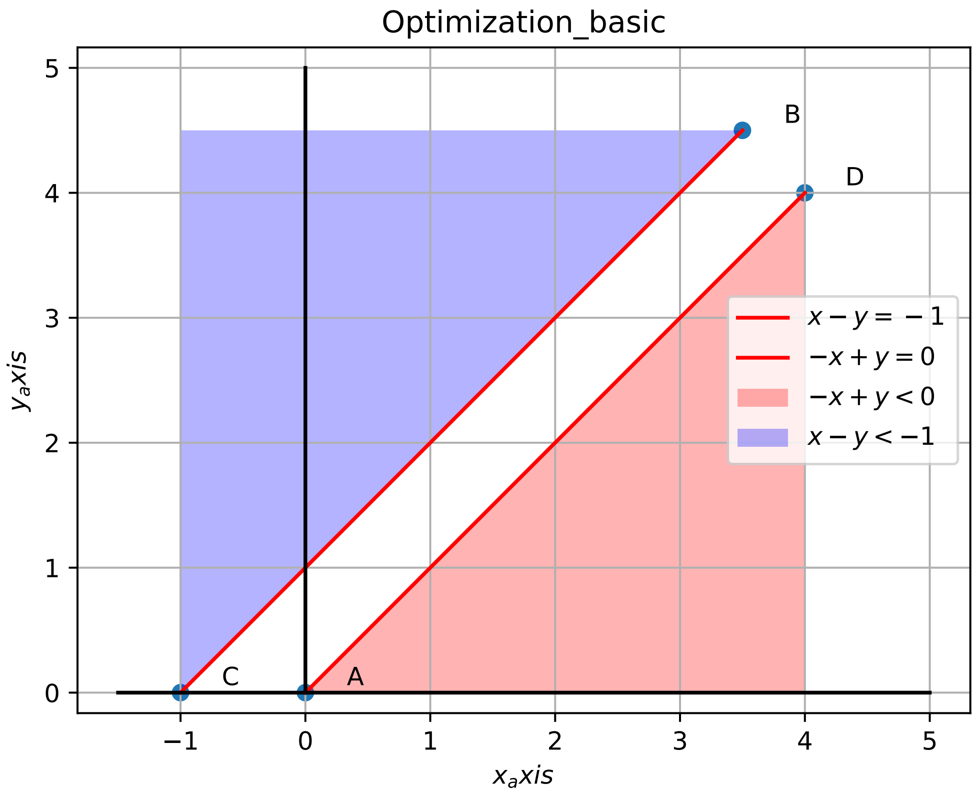


Figure 4.1.10.1:

verified from cvxpy by considering the following optimization problem.

$$z = \max_{x} \begin{pmatrix} 1 & 1 \end{pmatrix} \mathbf{x} \tag{4.1.10.5}$$

$$s.t. \quad \begin{pmatrix} 1 & -1 \\ -1 & 0 \\ 0 & -1 \\ -1 & 0 \end{pmatrix} x \preceq \begin{pmatrix} -1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \tag{4.1.10.6}$$

4.2. Applications

4.2.1 Reshma wishes to mix two types of food P and Q in such a way that the vitamin contents of the mixture contain at least 8 units of vitamin A and 11 units of vitamin B. Food P costs Rs 60/kg and Food Q costs Rs 80/kg. Food P contains 3 units/kg of Vitamin A and 5 units / kg of Vitamin B while food Q contains 4 units/kg of Vitamin A and 2 units/kg of vitamin B. Determine the minimum cost of the mixture.

Solution: Let the mixture contain x kg of food and y kg of food. The given information can be compiled in a table as and which can be expressed in vector form as

	Vitamin A(units/kg)	Vitamin B(units/kg)	Cost(Rs/kg)
Food P	3	5	60
Food Q	4	2	80
Requirement(units/kg)	8	11	

Table 4.2.1.1:

$$P = \min_{\mathbf{x}} \begin{pmatrix} 60 & 80 \end{pmatrix} \mathbf{x} \tag{4.2.1.1}$$

$$\begin{pmatrix} 3 & 4 \\ 5 & 2 \end{pmatrix} \mathbf{x} \succeq \begin{pmatrix} 8 \\ 11 \end{pmatrix} \tag{4.2.1.2}$$

$$\mathbf{x} \succeq \mathbf{0} \tag{4.2.1.3}$$

Solving using cvxpy, we get

$$P_{min} = 159.99999999 \tag{4.2.1.4}$$

$$\mathbf{x} = \begin{pmatrix} 2.11436236 \\ 0.41422823 \end{pmatrix} \tag{4.2.1.5}$$

4.2.2 One kind of cake requires 200g of flour and 25g of fat, and another kind of cake requires 100g of flour and 50g of fat. Find the maximum number of cakes which can be made from 5kg of flour and 1 kg of fat assuming that there is no shortage of the other ingredients used in making the cakes.

Solution: Let x, y be the number of cakes of first kind and second kind that can be made from the given amount of floor and fat respectively. From the given information,

Kind of cake	**No. of cakes**	**Flour (in gm)**	**Fat (in gm)**
$Cake_1$	x	200	25
$Cake_2$	y	100	50

Table 4.2.2.1:

$$200x + 100y \leq 5000 \tag{4.2.2.1}$$

$$100x + 50y \leq 1000 \tag{4.2.2.2}$$

Let P be the maximum number of cakes that can be made from the given amount of flour and fat. The problem can be formulated as

$$P = \max_{\mathbf{x}} \begin{pmatrix} 1 & 1 \end{pmatrix} \mathbf{x} \tag{4.2.2.3}$$

$$\begin{pmatrix} 200 & 100 \\ 100 & 50 \end{pmatrix} \mathbf{x} \leq \begin{pmatrix} 5000 \\ 1000 \end{pmatrix} \tag{4.2.2.4}$$

$$\mathbf{x} \geq \mathbf{0} \tag{4.2.2.5}$$

Solving the above equations using cvxpy, we get

$$P_{max} = 30, \mathbf{x} = \begin{pmatrix} 20 \\ 10 \end{pmatrix} \tag{4.2.2.6}$$

4.2.3 A factory makes tennis rackets and cricket bats. A tennis racket takes 1.5 hours of machine time and 3 hours of craftman's time in its making while a cricket bat takes 3 hour of machine time and 1 hour of craftman's time. In a day, the factory has the availability of not more than 42 hours of machine time and 24 hours of craftsman's time.

(a) What number of rackets and bats must be made if the factory is to work at full capacity?

(b) If the profit on a racket and on a bat is Rs 20 and Rs 10 respectively, find the maximum profit of the factory when it works at full capacity.

Solution: The given information is summarized in Table 4.2.3.1. From the given

Item	**Number**	**Machine hours**	**Craftman's hours**	**Profit**
Tennis Rackets	x	1.5	3	$Rs.20$
Cricket Bats	y	3	1	$Rs.10$
Maximum time available		42	24	

Table 4.2.3.1:

information, the optimization problem can be expressed as

$$Z = \max_{\mathbf{x}} \begin{pmatrix} 20 & 10 \end{pmatrix} \mathbf{x} \tag{4.2.3.1}$$

$$s.t. \quad \begin{pmatrix} 1 & 2 \\ 3 & 1 \end{pmatrix} \mathbf{x} \preceq \begin{pmatrix} 28 \\ 24 \end{pmatrix} \tag{4.2.3.2}$$

$$\mathbf{x} \succeq \mathbf{0} \tag{4.2.3.3}$$

From Fig. 4.2.3.1, the values at the corner points are obtained in Table 4.2.3.2.

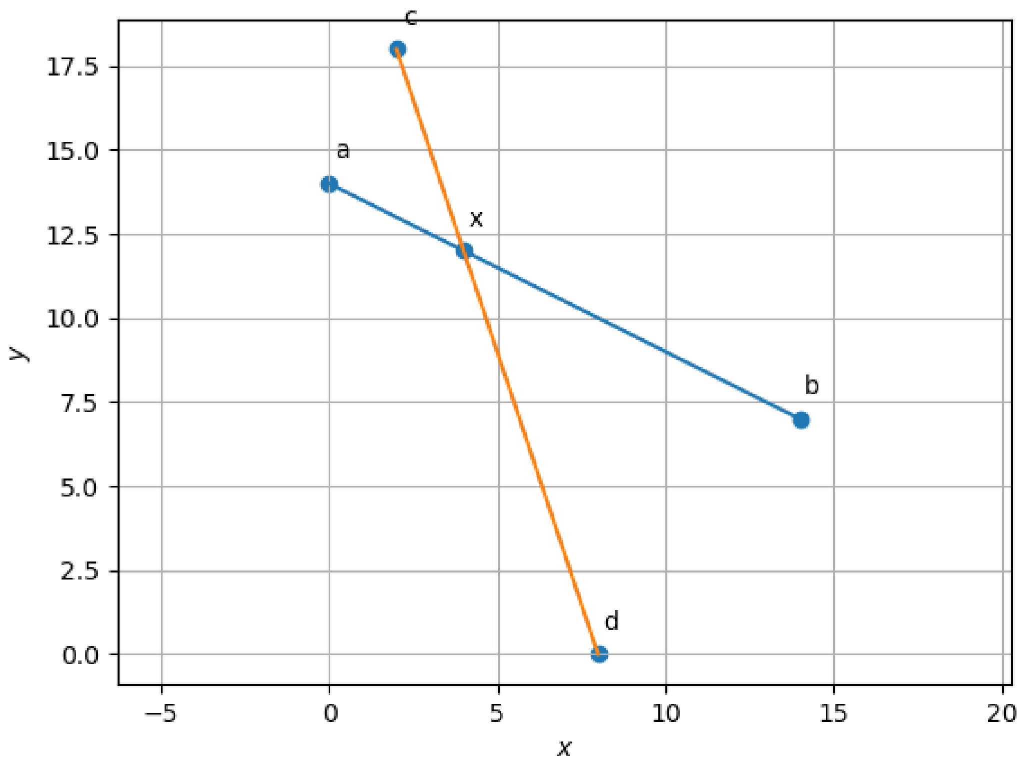


Figure 4.2.3.1:

Corner points	Value of Z
(0,14)	140
(4,12)	200
(8,0)	160

Table 4.2.3.2:

At full capacity,

$$\mathbf{x} = \begin{pmatrix} 4 \\ 12 \end{pmatrix}, \tag{4.2.3.4}$$

$$\implies Z = \begin{pmatrix} 20 & 10 \end{pmatrix} \begin{pmatrix} 4 \\ 12 \end{pmatrix} \tag{4.2.3.5}$$

This is verified using cvxpy.

4.2.4 A manufacturer produces nuts and bolts. It takes 1 hour of work on machine A and 3 hours on machine B to produce a package of nuts. It takes 3 hours on machine A and 1 hour on machine B to produce a package of bolts. He earns a profit of Rs17.50 per package on nuts and Rs 7.00 per package on bolts.

How many packages of each should be produced each day so as to maximise his profit, if he operates his machines for at the most 12 hours a day?.

Solution: Table 4.2.4.1 summarizes the given information. The optimization problem

Symbol	**Name**	**Machine A**	**Machine B**	**Profit**
x	nuts	1x	3x	17.5x
y	bolts	3y	1y	7y
Sum	x+y	x+3y	3x+y	17.5x+7y
t	time	12 h	12 h	

Table 4.2.4.1:

is formulated as

$$z = \begin{pmatrix} 17.5 & 7 \end{pmatrix} \mathbf{x} \tag{4.2.4.1}$$

$$\begin{pmatrix} 1 & 3 \\ 3 & 1 \end{pmatrix} \mathbf{x} \preceq \begin{pmatrix} 12 \\ 12 \end{pmatrix} \tag{4.2.4.2}$$

Fig. 4.2.4.1 represents the given constraints from which, the corner points of the feasible region intersected by above two lines are

$$\mathbf{0}, \mathbf{A} = \begin{pmatrix} 0 \\ 4 \end{pmatrix}, \mathbf{B} = \begin{pmatrix} 4 \\ 0 \end{pmatrix} \text{ and } \mathbf{C} = \begin{pmatrix} 3 \\ 3 \end{pmatrix} \tag{4.2.4.3}$$

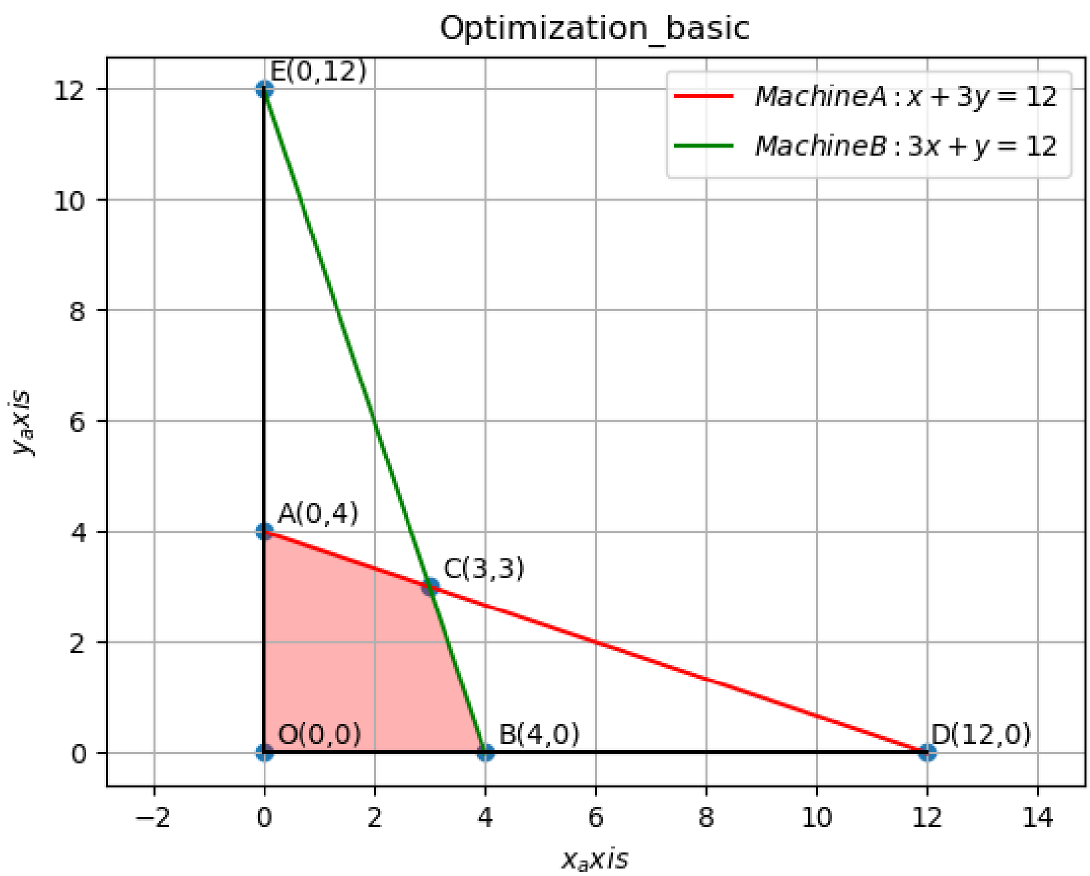


Figure 4.2.4.1:

The values of Z at these points are listed in Table 4.2.4.2. Thus, the maximum profit

Corner Point	**z=17.5x+7y**	**Remarks**
O=$\begin{pmatrix}0\\0\end{pmatrix}$	0	
A=$\begin{pmatrix}0\\4\end{pmatrix}$	28	
C=$\begin{pmatrix}3\\3\end{pmatrix}$	73.5	Maximum
B=$\begin{pmatrix}4\\0\end{pmatrix}$	70	

Table 4.2.4.2:

is

$$z = 73.50, \mathbf{x} = \begin{pmatrix} 3 \\ 3 \end{pmatrix}. \tag{4.2.4.4}$$

4.2.5 A factory manufactures two types of screws , A and B. Each type screw requires the use of two machines , an automatic and a hand operated. It takes 4 minutes on the automatic and 6 minutes on hand operated machines to manufacture a package of screws A,while it takes 6 minutes on automatic and 3 minutes on the hand operated machines to manufacture a package of screws B. Each machine is available for at the most 4hrs on any day. The manufacturer can sell a package of screws A at a profit of Rs. 7 and screws B at a profit of Rs. 10. Assuming that he can sell all the screws he manufactures, how many packages of each type should the factory owner produce in a day in order to maximise his profit? Determine the maximum profit.

Solution: The given information is summarized in Table 4.2.5.1 resulting in the

Item	**Number**	**Machine A**	**Machine B**	**Profit**
Screw A	x	4(min)	6(min)	7
SCREW B	y	6(min)	3(min)	10
Max.Time available		4 hrs	4 hrs	

Table 4.2.5.1:

following optimization problem.

$$z = \max_{\mathbf{x}} \begin{pmatrix} 7 & 10 \end{pmatrix} \begin{pmatrix} x \end{pmatrix} \tag{4.2.5.1}$$

$$s.t. \quad \begin{pmatrix} 2 & 3 \\ 2 & 1 \end{pmatrix} \preceq \begin{pmatrix} 120 \\ 80 \end{pmatrix} \tag{4.2.5.2}$$

$$\mathbf{x} \succeq \mathbf{0} \tag{4.2.5.3}$$

The maximum profit is obtained as

$$\begin{pmatrix} 2 & 3 & 120 \\ 2 & 1 & 80 \end{pmatrix} \longrightarrow R_2 - R_1 \begin{pmatrix} 2 & 3 & 120 \\ 0 & -2 & -40 \end{pmatrix} \longrightarrow R_2 * \frac{-1}{2} \tag{4.2.5.4}$$

$$\begin{pmatrix} 2 & 3 & 120 \\ 0 & 1 & 20 \end{pmatrix} \longrightarrow R_1 - 3R_2 \begin{pmatrix} 2 & 0 & 60 \\ 0 & 1 & 20 \end{pmatrix} \longrightarrow \frac{R_1}{2} \begin{pmatrix} 1 & 0 & 30 \\ 0 & 1 & 20 \end{pmatrix} \tag{4.2.5.5}$$

yielding the maximum profit

$$z = 410 \tag{4.2.5.6}$$

at

$$\mathbf{x} = \begin{pmatrix} 30 \\ 20 \end{pmatrix}. \tag{4.2.5.7}$$

This is verified by Table 4.2.5.2, where the corner points are obtained from Fig. 4.2.5.1.

Corner points	**Value of Z**
(0,40)	400
(30,20)	410
(40,0)	280

Table 4.2.5.2:

4.2.6

4.2.7

4.2.8 A merchant plans to sell two types of personal computers, a desktop model and a portable model that will cost Rs 25000 and Rs 40000 respectively. He estimates that the total monthly demand of computers will not exceed 250 units. Determine the

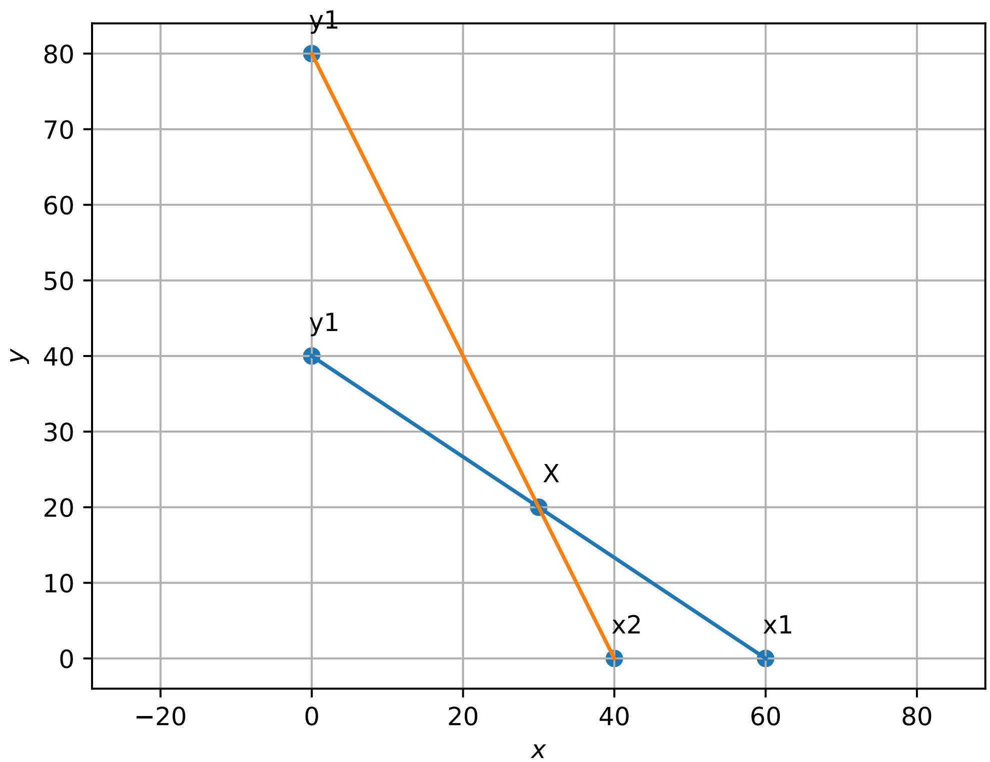


Figure 4.2.5.1:

number of units of each type of computers which the merchant should stock to get maximum profit if he does not want to invest more than Rs 70 lakhs and if his profit on the desktop model is Rs 4500 and on portable model is Rs 5000.

Solution: Table 4.2.8.1 summarizes the given information. The optimization problem

Item	**Number**	**Cost**	**Profit**
Desktop	x	25000	4500
Portable Computers	y	40000	5000
Max Investment		7000000	

Table 4.2.8.1:

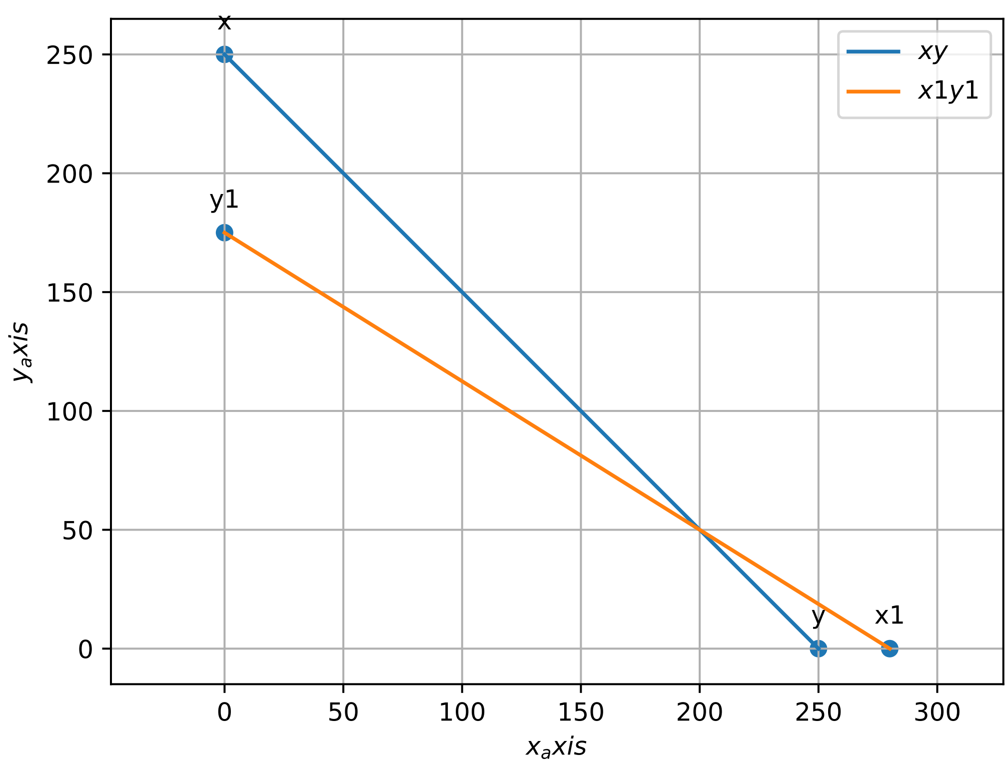


Figure 4.2.8.1:

can then be summarized as

$$Z = \begin{pmatrix} 4500 & 5000 \end{pmatrix} \mathbf{x} \tag{4.2.8.1}$$

$$s.t. \quad \begin{pmatrix} 1 & 1 \\ 5 & 8 \end{pmatrix} \mathbf{x} \preceq \begin{pmatrix} 250 \\ 1400 \end{pmatrix} \tag{4.2.8.2}$$

$$\mathbf{x} \succeq \mathbf{0} \tag{4.2.8.3}$$

From Fig. 4.2.8.1, the corner points are listed in Table 4.2.8.2 yielding the solution

$$Z = 1150000, \mathbf{x} = \begin{pmatrix} 200 \\ 50 \end{pmatrix} \tag{4.2.8.4}$$

Corner points	**Value of Z**
(250,0)	112500
(200,50)	1150000
(0,175)	875000

Table 4.2.8.2:

4.2.9

4.2.10 There are two types of fertilisers F_1 and F_2. F_1 consists of 10% Nitrogen and 6% Phosphoric acidand F_2 consists of 5% Nitrogen and 10% Phosphoric acid. After testing the soil conditions, a farmer finds that she needs atleast 14 kg of nitrogen and 14 kg of phosphoric acid for her crop. If F_1 costs Rs 6/kg and F_2 costs Rs 5/kg, determine how much of each type of fertiliser should be used so that nutrient requirements are met at a minimum cost. What is the minimum cost?

Solution: The optimization problem can be framed from Table 4.2.10.1 as

Fertiliser	**Nitrogen**	**Phosphoric Acid**
F_1	10%	6%
F_2	5%	10%
Total	14 kg	14 kg

Table 4.2.10.1:

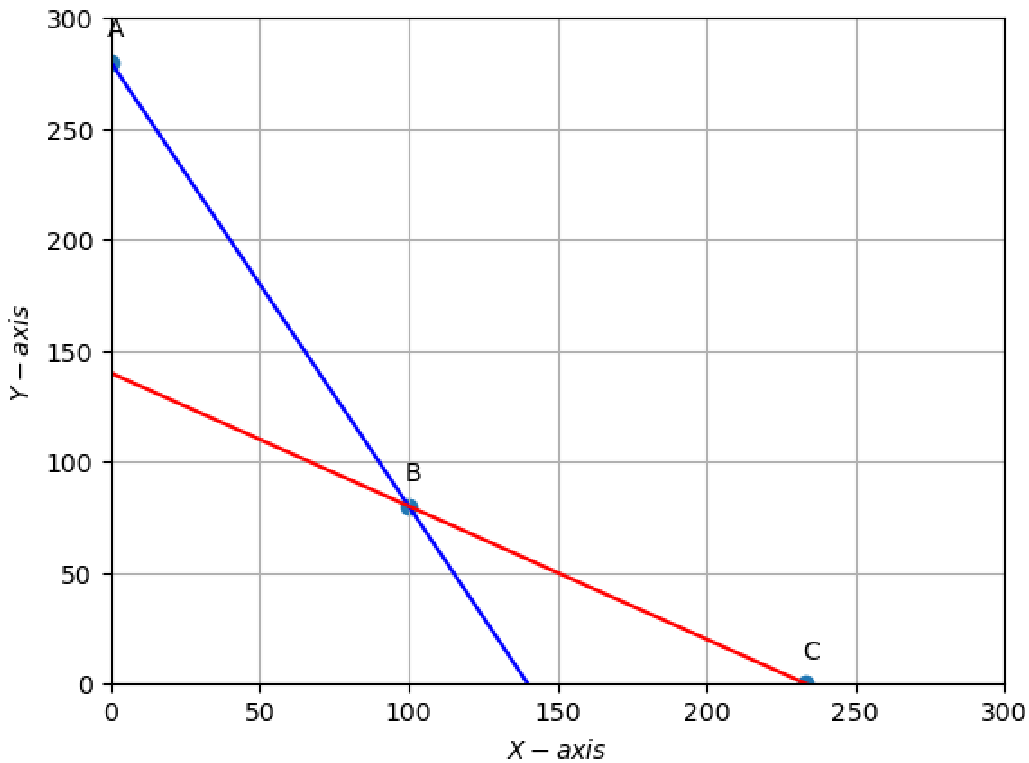


Figure 4.2.10.1:

$$Z = \min_{\mathbf{x}} \begin{pmatrix} 6 & 5 \end{pmatrix} \mathbf{x} \tag{4.2.10.1}$$

$$\begin{pmatrix} 2 & 1 \\ 3 & 5 \end{pmatrix} \mathbf{x} \preceq \begin{pmatrix} 280 \\ 700 \end{pmatrix} \tag{4.2.10.2}$$

$$\mathbf{x} \succeq \mathbf{0} \tag{4.2.10.3}$$

yielding

$$Z_{min} = Rs.1000, \mathbf{x} = \begin{pmatrix} 100 \\ 80 \end{pmatrix} \tag{4.2.10.4}$$

4.3. Additional

4.3.1 A dietician has to develop a special diet using two foods P and Q. Each packet (containing 30 g) of food P contains 12 units of calcium, 4 units of iron, 6 units of cholesterol and 6 units of vitamin A. Each packet of the same quantity of food Q contains 3 units of calcium, 20 units of iron, 4 units of cholesterol and 3 units of vitamin A. The diet requires atleast 240 units of calcium, atleast 460 units of iron and at most 300 units of cholesterol. How many packets of each food should be used to maximise the amount of vitamin A in the diet? What is the maximum amount of vitamin A in the diet?

4.3.2 A farmer mixes two brands P and Q of cattle feed. Brand P, costing Rs 250 per bag, contains 3 units of nutritional element A, 2.5 units of element B and 2 units of element C. Brand Q costing Rs 200 per bag contains 1.5 units of nutritional element A, 11.25 units of element B, and 3 units of element C. The minimum requirements of nutrients A, B and C are 18 units, 45 units and 24 units respectively. Determine the number of bags of each brand which should be mixed in order to produce a mixture having a minimum cost per bag? What is the minimum cost of the mixture per bag?

4.3.3 A dietician wishes to mix together two kinds of food X and Y in such a way that the mixture contains at least 10 units of vitamin A, 12 units of vitamin B and 8 units of vitamin C. The vitamin contents of one kg food is given below:

Food	Vitamin A	Vitamin B	Vitamin C
X	1	2	3
Y	2	2	1

Table 4.3.3.1:

One kg of food X costs Rs 16 and one kg of food Y costs Rs 20. Find the least cost of the mixture which will produce the required diet?

4.3.4 A manufacturer makes two types of toys A and B. Three machines are needed for this purpose and the time (in minutes) required for each toy on the machines is given below:

Types of Toys	Machines		
	I	II	III
A	12	18	6
B	6	18	9

Table 4.3.4.1:

Each machine is available for a maximum of 6 hours per day. If the profit on each toy of type A is Rs 7.50 and that on each toy of type B is Rs 5, show that 15 toys of type A and 30 of type B should be manufactured in a day to get maximum profit.

Solution: The given information can be framed as the optimization problem

$$Z = \max_{\mathbf{x}} \begin{pmatrix} 7.50 & 5 \end{pmatrix} \mathbf{x} \tag{4.3.4.1}$$

$$\begin{pmatrix} 2 & 1 \\ 3 & 0 \\ 2 & 3 \end{pmatrix} \mathbf{x} \preceq \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} \tag{4.3.4.2}$$

$$\mathbf{x} \succeq \mathbf{0} \tag{4.3.4.3}$$

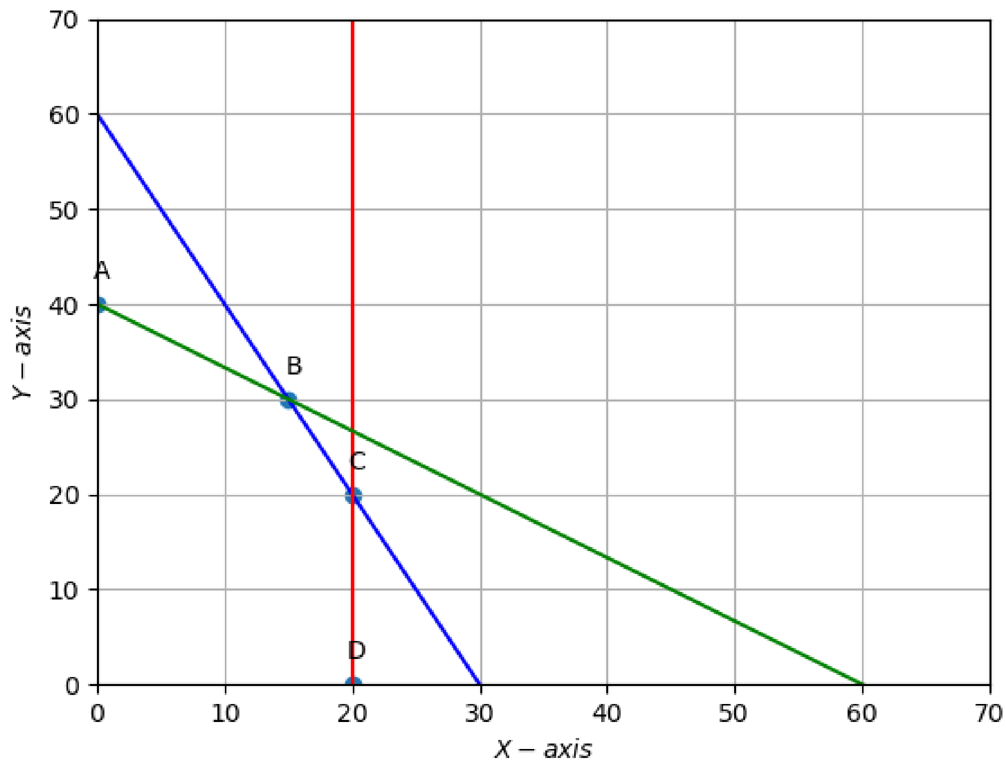


Figure 4.3.4.1:

Solving the above equations using cvxpy, we obtain

$$Z_{max} = Rs.262.50, \mathbf{x} = \begin{pmatrix} 15 \\ 30 \end{pmatrix} \tag{4.3.4.4}$$

4.3.5 An aeroplane can carry a maximum of 200 passengers. A profit of Rs 1000 is made on each executive class ticket and a profit of Rs 600 is made on each economy class ticket. The airline reserves at least 20 seats for executive class. However, at least 4 times as many passengers prefer to travel by economy class than by the executive class. Determine how many tickets of each type must be sold in order to maximise the profit for the airline. What is the maximum profit?

Solution: Let P be the maximum number of tickets of each type must be sold in order to maximise the profit for the airline . The problem can be formulated as

$$P = \max_{\mathbf{x}} \begin{pmatrix} 1 & 1 \end{pmatrix} \mathbf{x} \tag{4.3.5.1}$$

$$s.t. \quad \begin{pmatrix} -1 & -1 \\ 1 & 0 \\ -4 & 1 \end{pmatrix} \mathbf{x} \succeq \begin{pmatrix} 200 \\ 20 \\ 0 \end{pmatrix} \tag{4.3.5.2}$$

yielding

$$P_{max} = 136000, \mathbf{x} = \begin{pmatrix} 40 \\ 160 \end{pmatrix} \tag{4.3.5.3}$$

4.3.6 Two godowns A and B have grain capacity of 100 quintals and 50 quintals respectively. They supply to 3 ration shops, D, E and F whose requirements are 60, 50 and 40 quintals respectively. The cost of transportation per quintal from the godowns to the shops are given in the following table:

XUZ		
From/To	A	B
D	6	4
E	3	2
F	2.5	3

Table 4.3.6.1:

How should the supplies be transported in order that the transportation cost is min-

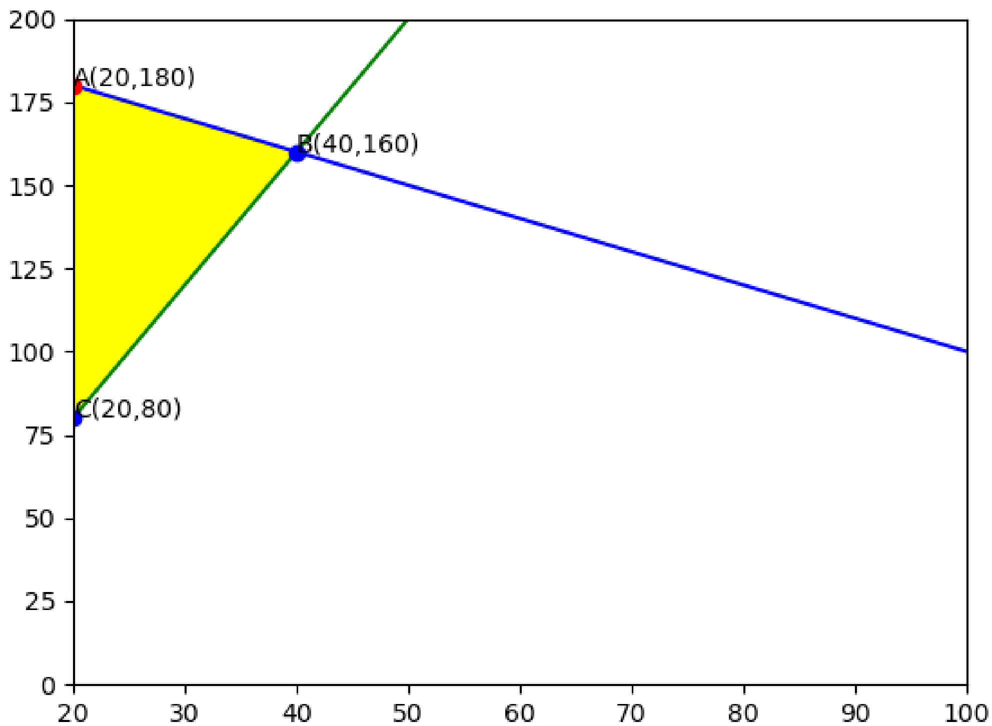


Figure 4.3.5.1:

imum? What is the minimum cost?

Solution: Let's assume that

(a) A supplies x quintals grain to ration shop D.

(b) A supplies y quintals grain to ration shop E.

(c) A will supply remaining grains 100-x-y quintals to F.

(d) B will supply 60-x quintals grain to ration shop D.

(e) B will supply 50-y quintals grain to ration shop E.

(f) B will supply x+y-60 quintals grain to ration shop F.

Total transportation cost is given by :

$$P = 2.5x + 1.5y + 410 \tag{4.3.6.1}$$

Now, Since godown A can supply maximum 60 quintals to ration shop D and 50 quintals to ration shop E and have maximum 100 quintals capacity to supply.

Also, if godown A supplies all 40 quintals to ration shop F, then remaining 60 quintals will be supplied to ration shop D and E and x and y is amount of grains. It can never be negative. This leads to the following conditions

$$x + y \leq 100 \tag{4.3.6.2}$$

$$x \leq 60 \tag{4.3.6.3}$$

$$y \leq 50 \tag{4.3.6.4}$$

$$-x - y \leq -60 \tag{4.3.6.5}$$

$$x \geq 0 \tag{4.3.6.6}$$

$$y \geq 0 \tag{4.3.6.7}$$

The optimization problem can then be expressed as

$$P = \max_{\mathbf{x}} \begin{pmatrix} 2.5 & 1.5 \end{pmatrix} \mathbf{x} + 410 \tag{4.3.6.8}$$

$$s.t. \quad \begin{pmatrix} 1 & 1 \\ -1 & -1 \\ -1 & 0 \\ 0 & -1 \end{pmatrix} \mathbf{x} \preceq \begin{pmatrix} 100 \\ -60 \\ -60 \\ -50 \end{pmatrix} \tag{4.3.6.9}$$

yielding

$$P = 510, \mathbf{x} = \begin{pmatrix} 10 \\ 50 \end{pmatrix} \tag{4.3.6.10}$$

Hence,

(a) The minimum transportation cost is : 510 /-

(b) A supplies 10 quintals grain to ration shop D.

(c) A supplies 50 quintals grain to ration shop E.

(d) A supplies 40 quintals grain to ration shop F.

(e) A supplies 50 quintals grain to ration shop D.

(f) A supplies 0 quintals grain to ration shop E.

(g) A supplies 0 quintals grain to ration shop F.

4.3.7 An oil company has two depots A and B with capacities of 7000 L and 4000 L respectively. The company is to supply oil to three petrol pumps, D, E and F whose requirements are 4500L, 3000L and 3500L respectively. The distances (in km) between the depots and the petrol pumps is given in the following table:

Distance in (km.)		
From / To	A	B
D	7	3
E	6	4
F	3	2

Table 4.3.7.1:

Assuming that the transportation cost of 10 litres of oil is Re 1 per km, how should the delivery be scheduled in order that the transportation cost is minimum? What is the minimum cost?

4.3.8 A fruit grower can use two types of fertilizer in his garden, brand P and brand Q. The amounts (in kg) of nitrogen, phosphoric acid, potash, and chlorine in a bag of each brand are given in the table. Tests indicate that the garden needs at least 240 kg of phosphoric acid, at least 270 kg of potash and at most 310 kg of chlorine

If the grower wants to minimise the amount of nitrogen added to the garden, how many bags of each brand should be used? What is the minimum amount of nitrogen added in the garden?

kg per bag		
	Brand P	Brand Q
Nitrogen	3	3.5
Phosphoric acid	1	2
Potash	3	1.5
Chlorine	1.5	2

Table 4.3.8.1:

Solution: The given information is summarized in Table 4.3.8.2.

Kg per bag	Brand P	Brand Q
Nitrogen	3	3.5
Phosphoric acid	1	2
Potash	3	1.5
Chlorine	1.5	2

Table 4.3.8.2:

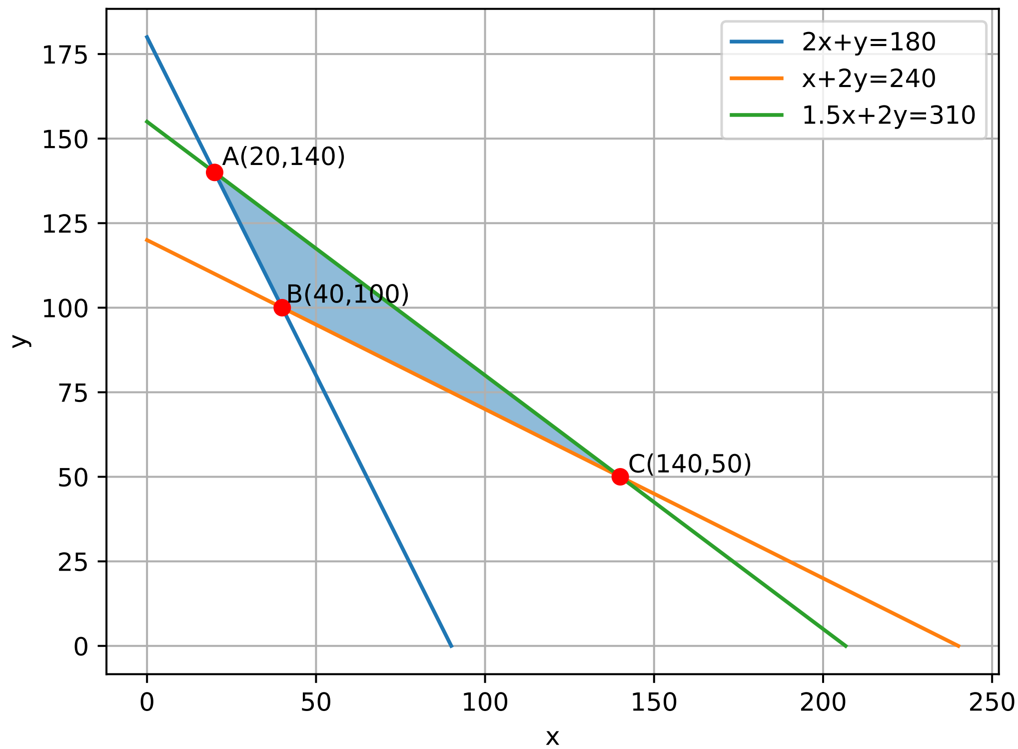


Figure 4.3.8.1:

The given problem can be expressed as

$$P = \min_{\mathbf{x}} \begin{pmatrix} 3 & 3.5 \end{pmatrix} \mathbf{x} \tag{4.3.8.1}$$

$$\begin{pmatrix} 2 & 1 \\ 1 & 2 \\ -1.5 & -2 \\ 1 & 0 \\ 0 & 1 \end{pmatrix} \mathbf{x} = \begin{pmatrix} 180 \\ 240 \\ -310 \\ 0 \\ 0 \end{pmatrix} \tag{4.3.8.2}$$

$$\mathbf{x} \succeq \mathbf{0} \tag{4.3.8.3}$$

yielding

$$P_{min} = 470, \mathbf{x} = \begin{pmatrix} 40 \\ 100 \end{pmatrix} \tag{4.3.8.4}$$

This can be verified using Fig. 4.3.8.1.

4.3.9 Refer to Question 4.3.8.1 If the grower wants to maximise the amount of nitrogen added to the garden, how many bags of each brand should be added? What is the maximum amount of nitrogen added?

4.3.10 A toy company manufactures two types of dolls, A and B. Market research and available resources have indicated that the combined production level should not exceed 1200 dolls per week and the demand for dolls of type B is at most half of that for dolls of type A. Further, the production level of dolls of type A can exceed three times the production of dolls of other type by at most 600 units. If the company makes profit of Rs 12 and Rs 16 per doll respectively on dolls A and B, how many of each should be produced weekly in order to maximise the profit?

Chapter 5

Miscellaneous

5.1. Examples

5.1.1 Find the maximum and minimum values of

(a) $f(x) = (2x-1)^2 + 3$ **Solution:** The given function has a minimum value as shown in Figure 5.1.1.1.

$$f'(x) = 8x - 4 \tag{5.1.1.1}$$

The minimum value of the function is calculated using Gradient Descent method as below

$$x_{n+1} = x_n - \alpha \nabla f(x_n) \tag{5.1.1.2}$$

Choosing

i. $\alpha = 0.001$

ii. precision = 0.0000001

iii. n = 10000000

iv. x_0 = -5

$$x_{min} = \frac{1}{2}, f(x)_{min} = 3 \tag{5.1.1.3}$$

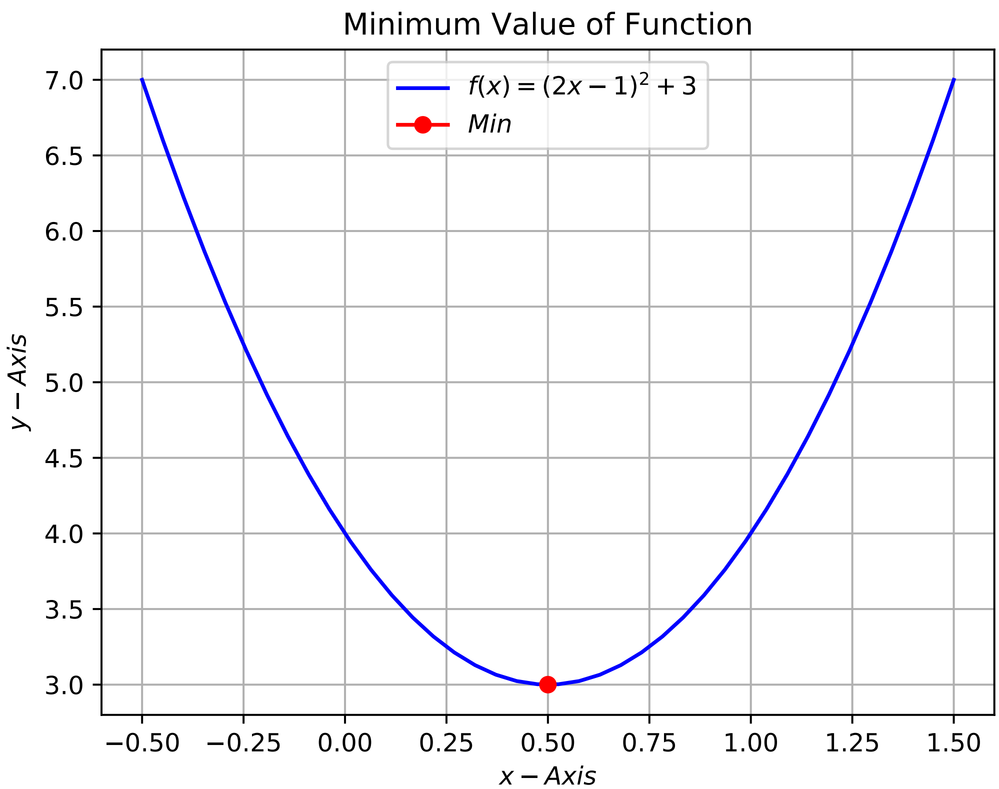


Figure 5.1.1.1:

5.1.2 At what points in the interval $(0,2\pi)$ does the function $\sin 2x$ attain its maximum value.

5.1.3 A square piece of tin of side 18 cm is to be made into a box without top, by cutting a square from each corner and folding up the flaps to form the box. What should be the side of the square to be cut off so that the volume of the box is the maximum possible.

Solution: Let the given side of tin be a.

$$a = 18cm \tag{5.1.3.1}$$

Lets cut a square of side x from each corner then the box formed folding up the flaps has dimensions as

$$l = a - 2x,\ b = a - 2x,\ h = x \tag{5.1.3.2}$$

The length, breadth, height are positive. These give the constraints on x

$$a - 2x > 0,\ x > 0 \tag{5.1.3.3}$$

$$\implies 0 < x < \frac{a}{2} \tag{5.1.3.4}$$

Volume of the box is given by

$$V\left(x\right) = x(a - 2x)^2 \tag{5.1.3.5}$$

(a) We now check the convexity of the function $V\left(x\right)$ under the constraints given by (5.1.3.4)

$$V'\left(x\right) = (a - 2x)(a - 6x) \tag{5.1.3.6}$$

$$V''\left(x\right) = 8(3x - a) \tag{5.1.3.7}$$

For $x > \frac{a}{3}$ the function $V\left(x\right)$ is convex. The function convexity is changing under the constraints. So the given problem cannot be expressed as a Convex optimization problem.

(b) The problem can be solved using the Gradient Descent Algorithm. It can be modified as

$$\min_{x} f(x) = -V(x) \tag{5.1.3.8}$$

Here we have,

$$f(x) = -x(a-2x)^2 \tag{5.1.3.9}$$

$$f'(x) = -(a-2x)(a-6x) \tag{5.1.3.10}$$

A numerical solution for (5.1.3.8) is obtained as

$$\lambda_{n+1} = \lambda_n - \alpha f'(x) \tag{5.1.3.11}$$

where λ_0 is an inital guess and μ is a variable parameter. These parameters decide how fast the algorithm converges. By taking the parameters as listed in the below table

Parameter	**Description**	**Value**
λ_0	Initial guess	8.5
α	Variable parameter	0.01
N	Number of iterations	10000
ϵ	Tolerance in λ	10^{-6}

Table 5.1.3.1:

The value of x obtained is

$$x = 3 \tag{5.1.3.12}$$

From equation (5.1.3.8), the minimum value of $f(x)$ is equivalent to maximum value of $V(x)$. Hence the maximum value of $V(x)$ occurs at $x = 3$.

$$V(3) = 432 \tag{5.1.3.13}$$

5.1.4 Find the absolute maximum and minimum values of the function f given by

$$f(x) = \cos^2 x + \sin x, \quad x \in [0, \pi] \tag{5.1.4.1}$$

5.1.5 Find the maximum profit that a company can make if the profit function is given by

$$f(x) = 41 - 72x + 18x^2 \tag{5.1.5.1}$$

Solution: Considering

$$\lambda\left(41 - 72x_1 - 18x_1^2\right) + (1-\lambda)\left(41 - 72x_2 - 18x_2^2\right) \geq \tag{5.1.5.2}$$

$$41 - 72\left(\lambda x_1 + (1-\lambda)x_2\right) - 18\left(\lambda x_1 + (1-\lambda)x_2\right)^2, \tag{5.1.5.3}$$

we obtain

$$18x_1^2\left(\lambda^2-\lambda\right)+18x_2^2\left(\lambda^2-\lambda\right)+36x_1x_2\left(\lambda^2-\lambda\right)\geq 0 \tag{5.1.5.4}$$

$$x_1^2\left(\lambda^2-\lambda\right)+x_2^2\left(\lambda^2-\lambda\right)+2x_1x_2\left(\lambda^2-\lambda\right)\geq 0 \tag{5.1.5.5}$$

$$-\lambda\left(1-\lambda\right)\left(x_1-x_2\right)^2\geq 0 \tag{5.1.5.6}$$

$$\implies \lambda\left(1-\lambda\right)\left(x_1-x_2\right)^2\leq 0 \tag{5.1.5.7}$$

which is false for all $\lambda \in (0,1)$. Hence the given function $f(x)$ is concave. Using the gradient ascent method,

$$x_n = x_{n-1} + \mu\frac{df(x)}{dx} \tag{5.1.5.8}$$

Since

$$\frac{df(x)}{dx} = -72 - 36x, \tag{5.1.5.9}$$

substituting (5.1.5.9) in (5.1.5.8),

$$x_n = x_{n-1} + \mu(-72 - 36x_{n-1}) \tag{5.1.5.10}$$

Choosing

$$x_0 = 1, \alpha = 0.001, precision = 0.00000001, \tag{5.1.5.11}$$

$$f_{max} \approx 113, x_{max} \approx -2.0, \tag{5.1.5.12}$$

which is verified in Fig. 5.1.5.1.

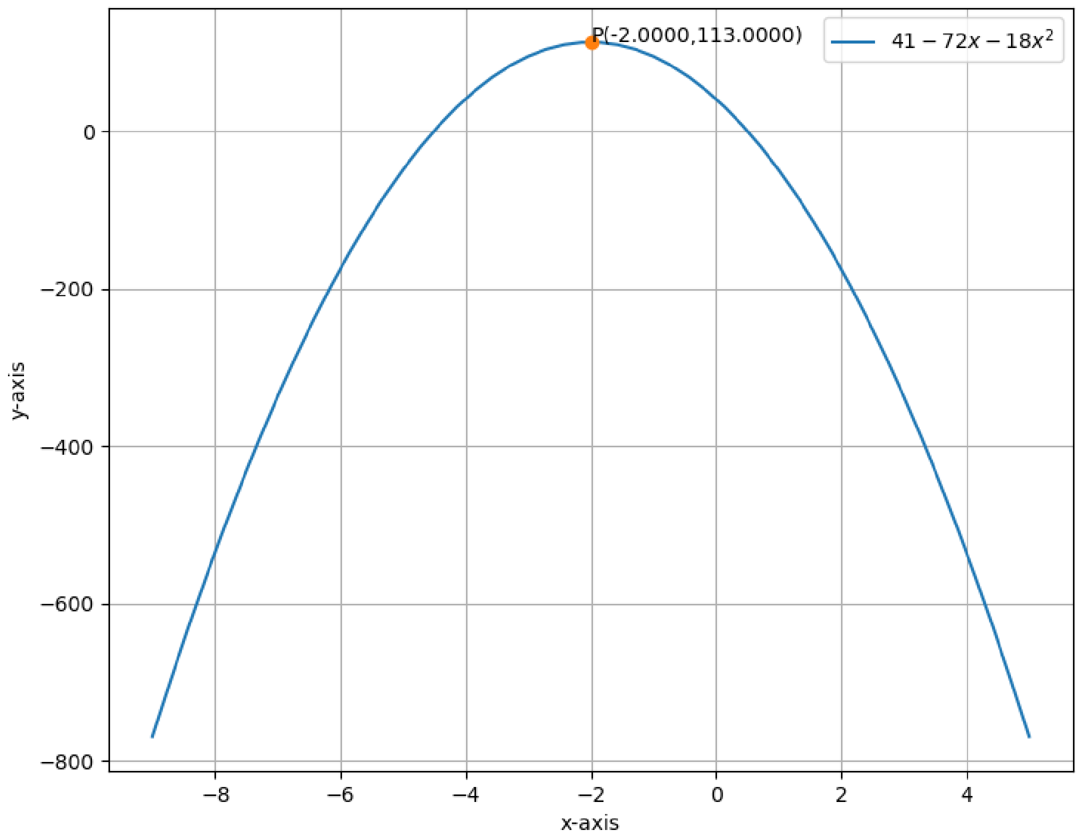


Figure 5.1.5.1:

5.1.6 Find both the maximum value and the minimum value of

$$f(x) = 3x^4 - 8x^3 + 12x^2 - 48x + 25 = 0 \qquad x \in (0, 3) \tag{5.1.6.1}$$

Solution:

$$\frac{df(x)}{dx} = 12x^3 - 24x^2 + 24x - 48 \tag{5.1.6.2}$$

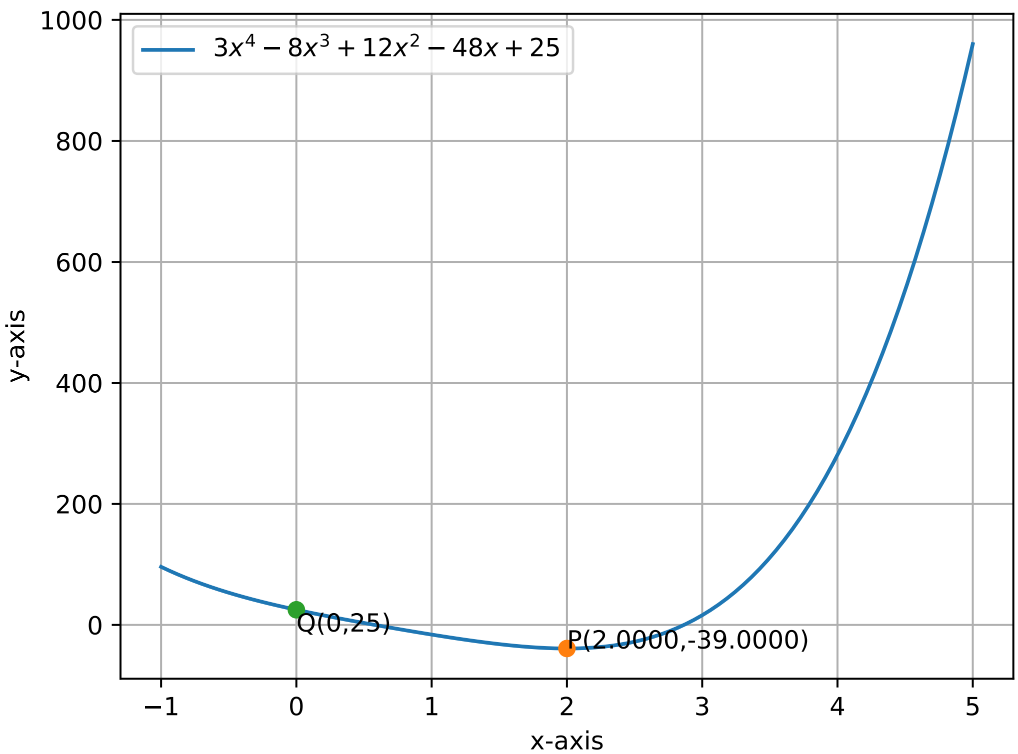


Figure 5.1.6.1:

The minimum can be found using

$$x_{n+1} = x_n - \alpha \frac{df(x)}{dx} \tag{5.1.6.3}$$

$$= x_n - \alpha(12x_n^3 - 24x_n^2 + 24x_n - 48) \tag{5.1.6.4}$$

where

(a) $\alpha = 0.001$

(b) x_{n+1} is current value

(c) x_n is previous value

(d) precession = 0.00000001

(e) maximum iterations = 100000000

as

$$f_{min} = -39 \tag{5.1.6.5}$$

$$x_{min} = 2 \tag{5.1.6.6}$$

5.1.7 At what points in the interval (0,2π) does the function $\sin 2x$ attain its maximum value.

Solution: Since

$$f(x) = \sin 2x, \tag{5.1.7.1}$$

$$f'(x) = 2\cos 2x \tag{5.1.7.2}$$

Using gradient ascent,

$$x_{n+1} = x_n + \alpha \nabla f(x_n) \tag{5.1.7.3}$$

$$= x_n + \alpha(2\cos 2x) \tag{5.1.7.4}$$

Choosing

$$x_0 = 0.5, \alpha = 0.001, precision = 0.00000001, \tag{5.1.7.5}$$

$$f_{max} = 1.0000, x_{max} = 0.7854. \tag{5.1.7.6}$$

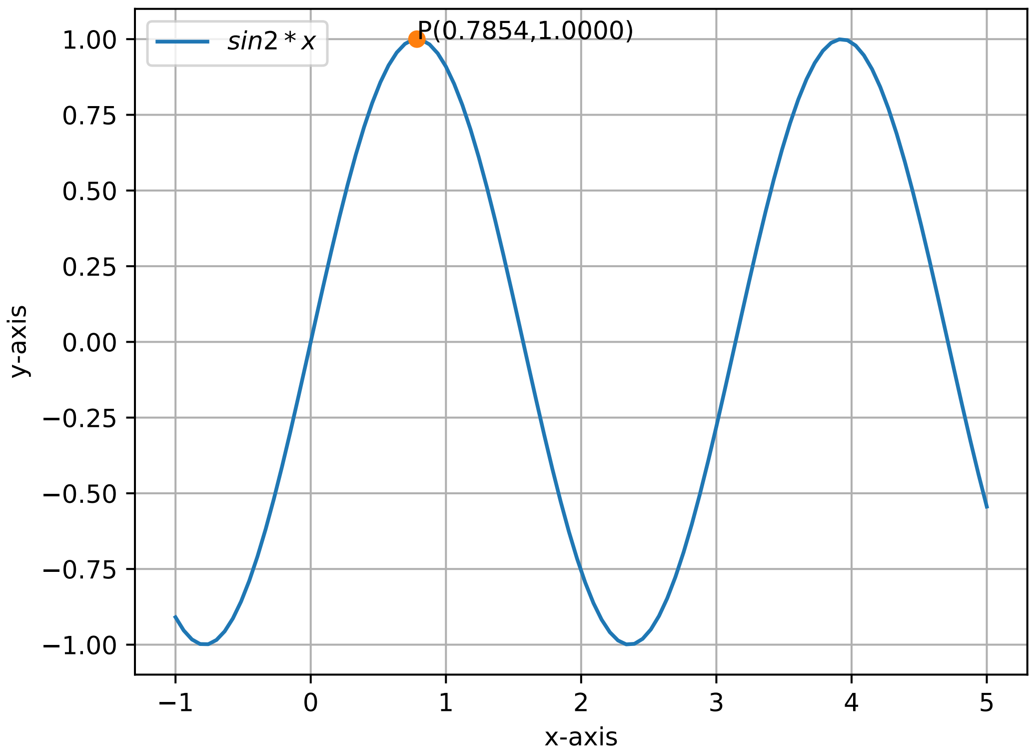


Figure 5.1.7.1:

5.1.8 Find the maximum value of $2x^3–24x + 107$ in the interval [1, 3]. Find the maximum value of the same function in [–3, –1].

Solution: Using gradient ascent method,

$$x_n = x_{n-1} + \mu \frac{df(x)}{dx} \tag{5.1.8.1}$$

where

$$\frac{df(x)}{dx} = 6x^2 - 24 \tag{5.1.8.2}$$

yielding

$$x_n = x_{n-1} + \mu(6x^2 - 24_{n-1}) \tag{5.1.8.3}$$

Choosing

$$x_0 = 1, \mu = 0.001 \text{ and precision} = 0.00000001, \tag{5.1.8.4}$$

$$f_{max} \approx 139, x_{max} \approx -2.0 \tag{5.1.8.5}$$

5.1.9 It is given that at x=1, the function $x^4 - 62x^2 + ax + 9$ attains its maximum value, on the interval [0,2]. Find the value of a.

Solution: Differentiating the given function,

$$\nabla f(x) = 4x^3 - 124x + a \tag{5.1.9.1}$$

Since f attains its maximum value on the interval [0,2] at $x = 1$,

$$\nabla f(1) = 0 \implies a = 120 \tag{5.1.9.2}$$

Using gradient descent,

$$x_{n+1} = x_n + \alpha \nabla f(x_n) \tag{5.1.9.3}$$

$$= x_n + \alpha \left(4x_n^3 - 124x_n + 120\right) \tag{5.1.9.4}$$

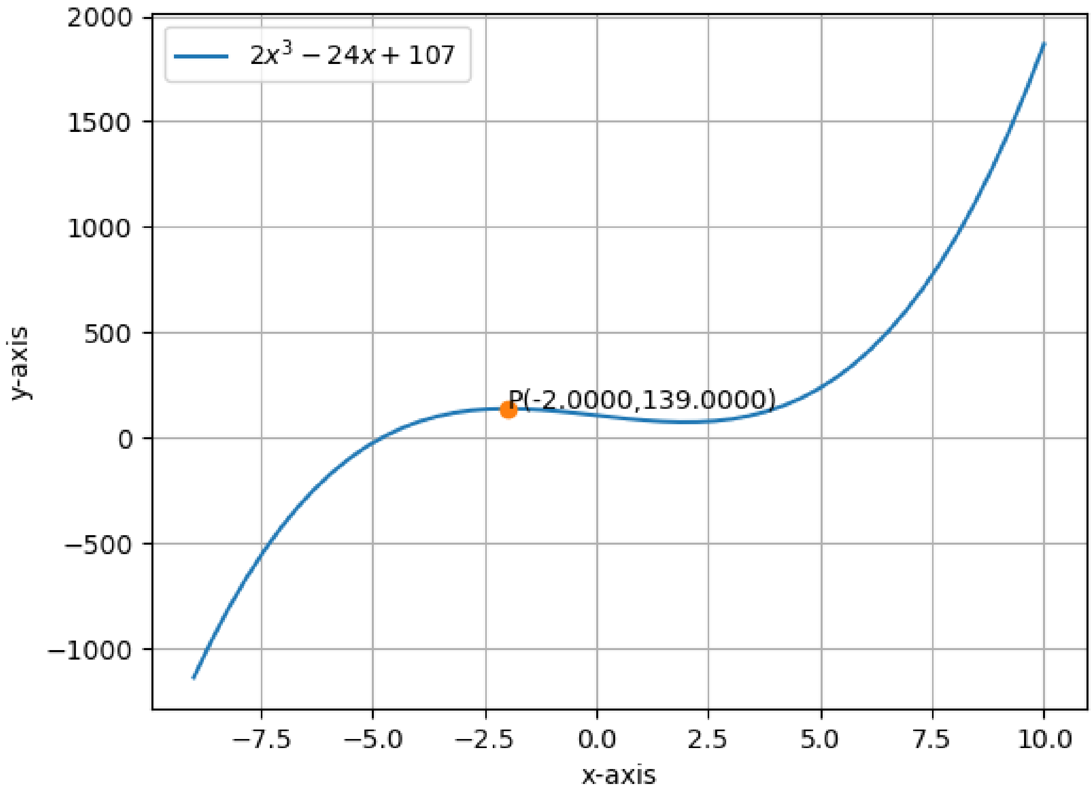


Figure 5.1.8.1:

and choosing

$$x_0 = 0.5, \alpha = 0.001 \text{ and precision} = 0.00000001, \tag{5.1.9.5}$$

$$f_{max} = 68, x_{max} = 1 \tag{5.1.9.6}$$

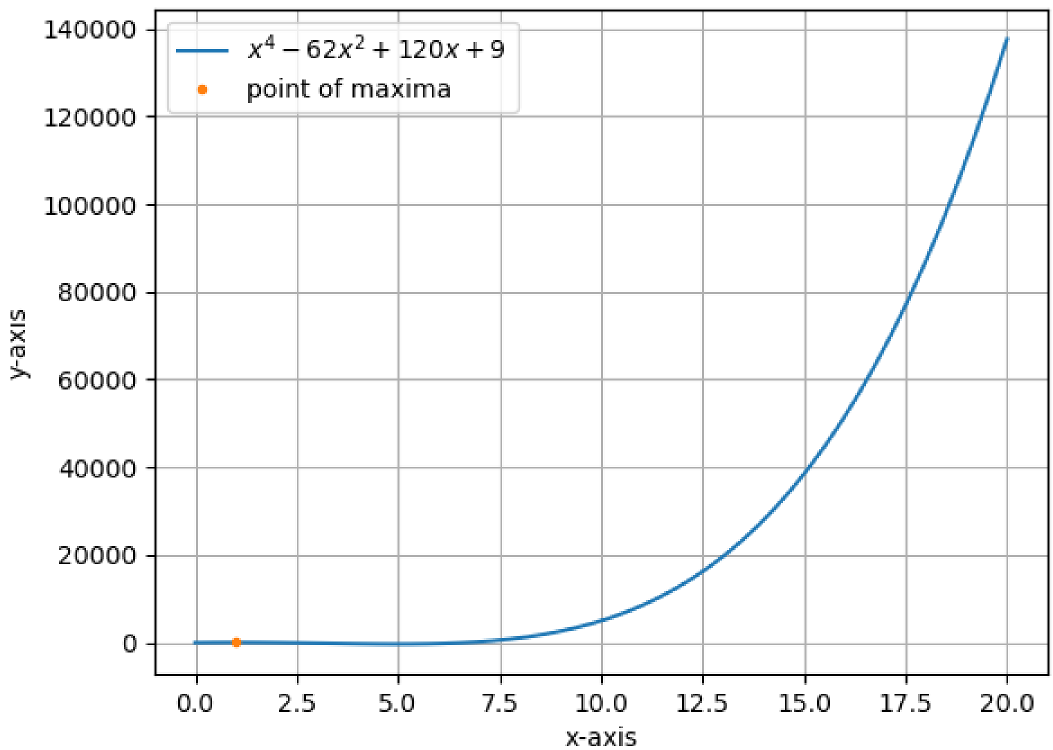


Figure 5.1.9.1:

5.1.10 Find the absolute maximum and minimum values of the function f given by

$$f(x) = \cos^2 x + \sin x, \quad x \in [0, \pi] \tag{5.1.10.1}$$

Solution: The derivative of the given function is

$$\nabla f(x) = \cos x - 2 \sin x \cos x \tag{5.1.10.2}$$

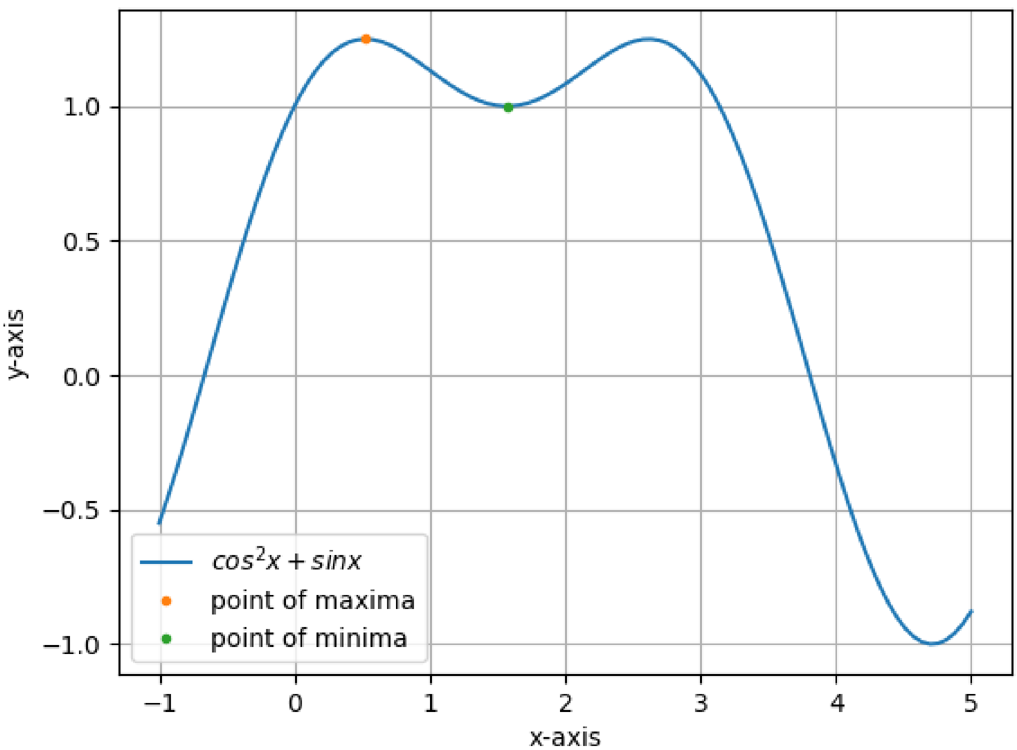


Figure 5.1.10.1:

The maxima is calculated by

$$x_{n+1} = x_n + \alpha \nabla f(x_n) \tag{5.1.10.3}$$

$$= x_n + \alpha \left(cosx_n - 2sinx_n cosx_n \right) \tag{5.1.10.4}$$

where

(a) $x_0 = 0.5$

(b) $\alpha = 0.001$

(c) precision = 0.00000001

yielding

$$f_{max} = 1.25, x_{max} = 0.52. \tag{5.1.10.5}$$

The minima is found by

$$x_{n+1} = x_n - \alpha \nabla f(x_n) \tag{5.1.10.6}$$

$$= x_n - \alpha \left(cosx_n - 2sinx_ncosx_n \right) \tag{5.1.10.7}$$

5.1.11 A square piece of tin of side 18 cm is to be made into a box without top, by cutting a square from each corner and folding up the flaps to form the box. What should be the side of the square to be cut off so that the volume of the box is the maximum possible.

Solution: Let the given side of tin be a.

$$a = 18cm \tag{5.1.11.1}$$

Lets cut a square of side x from each corner then the box formed folding up the flaps has dimensions as

$$l = a - 2x,\ b = a - 2x,\ h = x \tag{5.1.11.2}$$

The length, breadth, height are positive. These give the constraints on x

$$a - 2x > 0,\ x > 0 \tag{5.1.11.3}$$

$$\implies 0 < x < \frac{a}{2} \tag{5.1.11.4}$$

Volume of the box is given by

$$V(x) = x(a - 2x)^2 \tag{5.1.11.5}$$

The given problem can be expressed as a constrained optimization problem as

$$\max_{x} f(x) \triangleq V(x) \tag{5.1.11.6}$$

$$\text{s.t.} \quad g_1(x) \triangleq -x \leq 0 \tag{5.1.11.7}$$

$$g_2(x) \triangleq x \leq \frac{a}{2} \tag{5.1.11.8}$$

Define

$$L(x, \lambda, \mu) = f(x) + \lambda g_1(x) + \mu g_2(x) \tag{5.1.11.9}$$

Using the KKT conditions for finding the optimal point we have,

$$\implies \nabla L(x, \lambda, \mu) = 0 \tag{5.1.11.10}$$

$$\text{subject to} \quad \lambda g_1(x) = 0 \tag{5.1.11.11}$$

$$\mu g_2(x) = 0 \tag{5.1.11.12}$$

$$\lambda \geq 0, \mu \geq 0 \tag{5.1.11.13}$$

Here we get

$$\nabla f(x) = (a - 2x)(a - 6x) \tag{5.1.11.14}$$

$$\nabla g_1(x) = -1 \tag{5.1.11.15}$$

$$\nabla g_2(x) = 1 \tag{5.1.11.16}$$

By using the equations (5.1.11.14), (5.1.11.15), (5.1.11.16) we get

$$(a - 2x)(a - 6x) - \lambda + \mu = 0 \tag{5.1.11.17}$$

$$\lambda = 0 \tag{5.1.11.18}$$

$$\mu = 0 \tag{5.1.11.19}$$

Hence the value of x is given by,

$$x = \frac{a}{2}, \frac{a}{6} \tag{5.1.11.20}$$

For the value of $x = \frac{a}{2}$, is a boundary point for which the volume is zero. Hence the maximum volume is given by

$$\max V(x) = V\left(\frac{a}{6}\right) \tag{5.1.11.21}$$

$$= \frac{2}{27}a^3 \tag{5.1.11.22}$$

$$= 432 \tag{5.1.11.23}$$

5.1.12 The point on the curve

$$x^2 = 2y \tag{5.1.12.1}$$

which is nearest to the point $\mathbf{P} = \begin{pmatrix} 0 \\ 5 \end{pmatrix}$ is

(a) $\begin{pmatrix} 2\sqrt{2} \\ 4 \end{pmatrix}$

(b) $\begin{pmatrix} 2\sqrt{2} \\ 0 \end{pmatrix}$

(c) $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$

(d) $\begin{pmatrix} 2 \\ 2 \end{pmatrix}$

Solution: We need to find

$$\min_{\mathbf{x}} g(\mathbf{x}) = \|\mathbf{x} - \mathbf{P}\|^2 \tag{5.1.12.2}$$

$$\text{s.t. } h(\mathbf{x}) = \mathbf{x}^\top \mathbf{V}\mathbf{x} + 2\mathbf{u}^\top \mathbf{x} + f = 0 \tag{5.1.12.3}$$

where

$$\mathbf{V} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \ \mathbf{u} = \begin{pmatrix} 0 \\ -1 \end{pmatrix}, \ f = 0 \tag{5.1.12.4}$$

Suppose $\mathbf{x_1}$ and $\mathbf{x_2}$ satisfy $h(\mathbf{x}) = 0$. Then,

$$\mathbf{x_1}^\top \mathbf{V}\mathbf{x_1} + 2\mathbf{u}^\top \mathbf{x_1} + f = 0 \tag{5.1.12.5}$$

$$\mathbf{x_2}^\top \mathbf{V}\mathbf{x_2} + 2\mathbf{u}^\top \mathbf{x_2} + f = 0 \tag{5.1.12.6}$$

Then, for any $0 \leq \lambda \leq 1$, substituting

$$\mathbf{x} = \lambda \mathbf{x_1} + (1 - \lambda)\, \mathbf{x_2} \tag{5.1.12.7}$$

into (5.1.12.3), we get

$$h(\mathbf{x}) = \lambda (\lambda - 1) (\mathbf{x_1} - \mathbf{x_2})^\top \mathbf{V} (\mathbf{x_1} - \mathbf{x_2}) \neq 0 \tag{5.1.12.8}$$

since $\mathbf{x_1} - \mathbf{x_2}$ can be arbitrary. Hence, the optimization problem is nonconvex as the set of points on the parabola do not form a convex set. The constraints throw an error when *cvxpy* is used.

Chapter 6

Exemplar

1. At what point ,the slope of the curve $y = -x^3 + 3x^2 + 9x - 27$ is maximum? Also find the maximum slope.

2. prove that f (x)=sinx+$\sqrt{3}$ cosx has maximum value at $x = \frac{\pi}{6}$

Long Answer (L.A)

3. If the sum of the lengths of the hypotenuse and a side of a right angled triangle is given, show that the area of the triangle is maximum when the angle between them is $\frac{\pi}{3}$

4. Find the points of local maximal, local minima and the points of inflection of the function f (x)=$x^5 - 5x4 + 5x^2 - 1$. Also find the corresponding local maximam and local minimum values

5. A telephone company in a town has 500 subscribers on its and collects fixed charges of RS 300/- per subscriber per year. The company proposes to increase the annual subscription and it is belived that fog every increase of Re1/. one subsariber will discontinue the serice. find what increase will bring maximum profit?

6. If the straight line x casoc+y sno oc=P touches the curve $\frac{x^2}{a^2}+\frac{y^2}{b^2}=1$,then prove that $a^2\cos^2\alpha+b^2\sin^2\alpha=P^2$.

7. An open box with square base is to be made of a given quantity of card board of area c^2. show that the maiximum volume of the box is $\frac{c^23}{6\sqrt{3}}$cubic units.

8. Find the bimensions of the rectangle of perimeter 36cm which will sweep out a volume as large as possible, when revolved about one of its sides. Also find the maximum voiume.

9. lf the sum of the surface areas of cube and a sphere is constant, what is the ratio of on edge of the cube to the diameter of the sphere when the sum of their volumes is minimum?

10. AB is a diameter of a circle and c is any point on the circle show that the area of △ABC is maximum, when it is isosceles,

11. A metal box with a square base and vertical sides is to contain $1024cm^3$ The material for the top and bottom casts $RS5/cm^2$ and the material for the sides costs Rs2.50$/cm^2$.Find the least cost of the box.

12. The sum of the surface areas of a rectangular parallelopiped with sides x,2x and $\frac{x}{3}$ and a sphere is given to be constant. prove that the sum of their volumes is minimum, if x is equal to three times the radius of the sphere Also find the minimum value of the sum of their volumes

13. The least value of the function f (x)=ax+$\frac{b}{x}$(a>0,b>0,x>0) is ______

14. if x is real ,the minimum value of $x^2-8x+17$is

    (a) -1

(b) 0

(c) 1

(d) 2

15. The smallest value of the polynomial $x^3 - 18x^2 + 96x$ in [0,9] is

(a) 126

(b) 0

(c) 135

(d) 160

16. The function f (x)=$2x^3 - 3x^2 - 12x + 4$ has

(a) two points of local maximum

(b) two points of local minimum

(c) one maxima and one minima

(d) no marima or minima

17. The maximum volue of sin x casx is

(a) $\frac{1}{4}$

(b) $\frac{1}{2}$

(c) $\sqrt{2}$

(d) $2\sqrt{2}$

18. At x $=5\frac{\pi}{6}$ f (x)=2 sin3x+3 cosx

(a) maximum

(b) minimum

(c) zero

(d) neither maximum nor minimum

19. maximum slope of the curve y=$-x^3+3x^2+9x-27$ is

(a) 0

(b) 12

(c) 16

(d) 32

20. f (x)=x^2 has a stationary point at

(a) x=e

(b) x=$\frac{1}{e}$

(c) x=1

(d) x=$\sqrt{e}$

21. The maximum value of $[\frac{1}{x}]$ is

(a) e

(b) e^e

(c) $e^{\frac{1}{e}}$

(d) $\frac{1}{e}^{\frac{1}{e}}$

Appendix A

Convex Functions

A.1. Express the problem of finding the distance of the point $\mathbf{P} = \begin{pmatrix} 8 \\ 6 \end{pmatrix}$ from the line

$$L: \quad \begin{pmatrix} 1 & 1 \end{pmatrix} \mathbf{x} = 9 \tag{A.1.1}$$

as an optimization problem.

Solution: The given problem can be expressed as

$$\min_{\mathbf{x}} g(\mathbf{x}) = \left\|\mathbf{x} - \mathbf{P}\right\|^2 \tag{A.1.2}$$

$$\text{s.t.} \quad \mathbf{n}^\top \mathbf{x} = c \tag{A.1.3}$$

where

$$\mathbf{n} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, c = 9 \tag{A.1.4}$$

A.2. Explain Problem A.1 through a plot and find a graphical solution.

A.3. Solve (A.1.2) using cvxpy.

Solution: The following code yields

opt/codes/line_dist_cvx.py

$$\mathbf{x}_{\min} = \begin{pmatrix} 5.5 \\ 3.5 \end{pmatrix}, \tag{A.3.1}$$

$$g\left(\mathbf{x}_{\min}\right) = 3.53 \tag{A.3.2}$$

A.4. Convert (A.1.2) to an unconstrained optimization problem.

Solution: L in (A.1.1) can be expressed in terms of the direction vector $\mathbf{m}$ as

$$\mathbf{x} = \mathbf{A} + \lambda\mathbf{m}, \tag{A.4.1}$$

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

$$\mathbf{m}^T\mathbf{n} = 0 \tag{A.4.2}$$

Substituting (A.4.1) in (A.1.2), an unconstrained optimization problem

$$\min_{\lambda} f(\lambda) = \left\|\mathbf{A} + \lambda\mathbf{m} - \mathbf{P}\right\|^2 \tag{A.4.3}$$

is obtained.

A.5. Solve (A.4.3).

Solution:

$$f(\lambda) = (\lambda\mathbf{m} + \mathbf{A} - \mathbf{P})^T (\lambda\mathbf{m} + \mathbf{A} - \mathbf{P}) \tag{A.5.1}$$

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

$$+ \|\mathbf{A} - \mathbf{P}\|^2 \tag{A.5.2}$$

$$\because f^{(2)}\lambda = 2\|\mathbf{m}\|^2 > 0 \tag{A.5.3}$$

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

$$f^{(1)}(\lambda) = 2\lambda \|\mathbf{m}\|^2 + 2\mathbf{m}^T (\mathbf{A} - \mathbf{P}) = 0 \tag{A.5.4}$$

$$\implies \lambda_{\min} = -\frac{\mathbf{m}^T (\mathbf{A} - \mathbf{P})}{\|\mathbf{m}\|^2} \tag{A.5.5}$$

Choosing $\mathbf{A}$ such that

$$\mathbf{m}^T (\mathbf{A} - \mathbf{P}) = 0, \tag{A.5.6}$$

substituting in (A.5.5),

$$\lambda_{\min} = 0 \quad \text{and} \tag{A.5.7}$$

$$\mathbf{A} - \mathbf{P} = \mu\mathbf{n} \tag{A.5.8}$$

for some constant μ. (A.5.8) is a consequence of (A.4.2) and (A.5.6). Also, from

(A.5.8),

$$\mathbf{n}^T(\mathbf{A}-\mathbf{P}) = \mu \|\mathbf{n}\|^2 \tag{A.5.9}$$

$$\implies \mu = \frac{\mathbf{n}^T\mathbf{A}-\mathbf{n}^T\mathbf{P}}{\|\mathbf{n}\|^2} = \frac{c-\mathbf{n}^T\mathbf{P}}{\|\mathbf{n}\|^2} \tag{A.5.10}$$

from (A.1.3). Substituting $\lambda_{\min} = 0$ in (A.4.3),

$$\min_{\lambda} f(\lambda) = \|\mathbf{A}-\mathbf{P}\|^2 = \mu^2 \|\mathbf{n}\|^2 \tag{A.5.11}$$

upon substituting from (A.5.8). The distance between $\mathbf{P}$ and L is then obtained from (A.5.11) as

$$\|\mathbf{A}-\mathbf{P}\| = |\mu| \|\mathbf{n}\| \tag{A.5.12}$$

$$= \frac{\left|\mathbf{n}^T\mathbf{P}-c\right|}{\|\mathbf{n}\|} \tag{A.5.13}$$

after substituting for μ from (A.5.10). Using the corresponding values from Problem (A.1) in (A.5.13),

$$\min_{\lambda} f(\lambda) = 0.6 \tag{A.5.14}$$

A.6. The following python script plots

$$f(\lambda) = a\lambda^2 + b\lambda + d \tag{A.6.1}$$

for

$$a = \|\mathbf{m}\|^2 > 0 \tag{A.6.2}$$

$$b = \mathbf{m}^T(\mathbf{A} - \mathbf{P}) \tag{A.6.3}$$

$$c = \|\mathbf{A} - \mathbf{P}\|^2 \tag{A.6.4}$$

where $\mathbf{A}$ is the intercept of the line L in (A.1.1) on the x-axis and the points

$$\mathbf{U} = \begin{pmatrix} \lambda_1 \\ f(\lambda_1) \end{pmatrix}, \mathbf{V} = \begin{pmatrix} \lambda_2 \\ f(\lambda_2) \end{pmatrix} \tag{A.6.5}$$

$$\mathbf{X} = \begin{pmatrix} t\lambda_1 + (1-t)\lambda_2 \\ f[t\lambda_1 + (1-t)\lambda_2] \end{pmatrix}, \tag{A.6.6}$$

$$\mathbf{Y} = \begin{pmatrix} t\lambda_1 + (1-t)\lambda_2 \\ tf(\lambda_1) + (1-t)f(\lambda_2) \end{pmatrix} \tag{A.6.7}$$

for

$$\lambda_1 = -3, \lambda_2 = 4, t = 0.3 \tag{A.6.8}$$

in Fig. A.6.1. 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

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

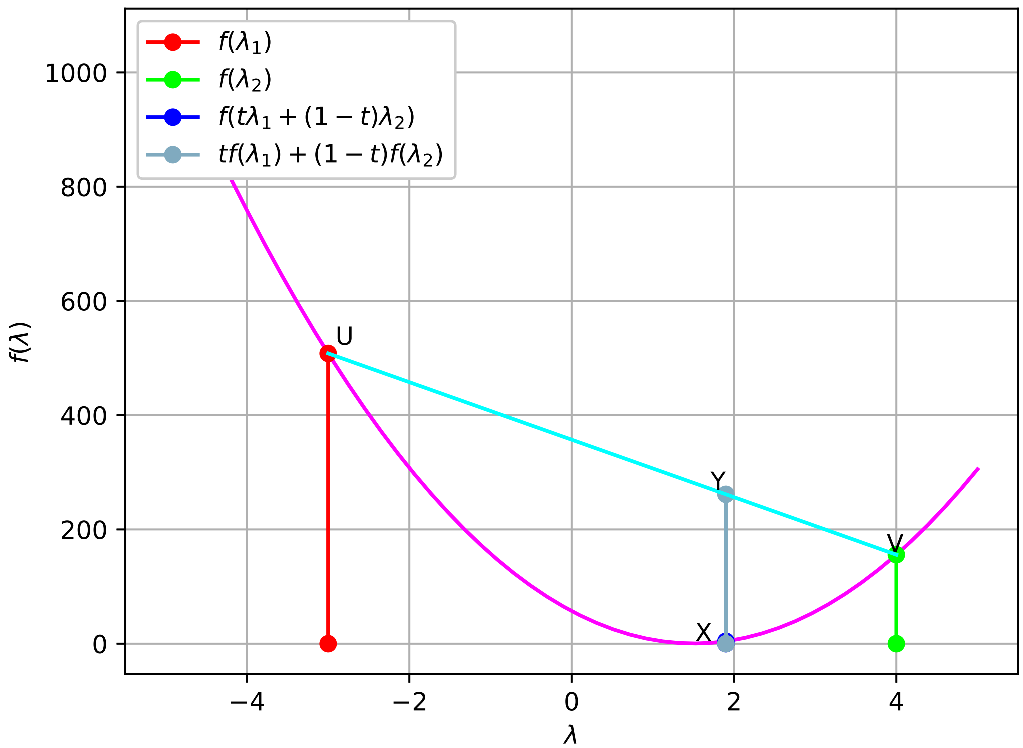


Figure A.6.1: $f(\lambda)$ versus λ

.

A.7. Show that

$$f\left[t\lambda_1 + (1-t)\,\lambda_2\right] \le tf\,(\lambda_1) + (1-t)\,f\,(\lambda_2) \tag{A.7.1}$$

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

A.8. Show that

$$\text{(A.7.1)} \quad \implies f^{(2)}(\lambda) > 0 \tag{A.8.1}$$

A.9. Show that a covex function has a unique minimum.

A.10. A single variable function f is said to be convex if

$$f\left[\mu x_1 + (1-\mu)\, x_2\right] \leq \mu f\left(x_1\right) + (1-\mu)\, f\left(x_2\right), \tag{A.10.1}$$

for $\quad 0 < \mu < 1$ and $x_1, x_2 \in \mathbb{R}$.

For a generic quadratic function $ax^2 + bc + c$, let us determine the sufficient condition for it to be convex. Let

$$f\left(x\right) = ax^2 + bx + c \tag{A.10.2}$$

Substituting LHS of inequality from (A.10.1) in (A.10.2)

$$\begin{aligned}
f\left[\mu x_1 + (1-\mu)\, x_2\right] &= f\left[x_2 + \mu\left(x_1 - x_2\right)\right] \\
&\implies a\left[x_2 + \mu\left(x_1 - x_2\right)\right]^2 b\left[x_2 + \mu\left(x_1 - x_2\right)\right] + c \\
&\quad\implies ax_2^2 + a\mu^2 x_1^2 + a\mu^2 x_2^2 - 2a\mu^2 x_1 x_2 \\
&\qquad + 2a\mu x_1 x_2 - 2a\mu x_2^2 + bx_2 + b\mu x_1 - b\mu x_2 + c
\end{aligned} \tag{A.10.3}$$

Substituting RHS of inequality from (A.10.1) in (A.10.2)

$$\begin{aligned}
\mu f\left(x_1\right) + (1-\mu)\, f\left(x_2\right) &= a\mu x_1^2 + b\mu x_1 + \mu c \\
&\quad + (1-\mu)\left(ax_2^2 + bx_2 + c\right) \\
&\implies a\mu x_1^2 + b\mu x_1 + ax_2^2 + bx_2 + c - a\mu x_2^2 - b\mu x_2
\end{aligned} \tag{A.10.4}$$

Combining (A.10.3) and (A.10.4) with inequality and simplifying

$$a\mu^2x_1^2 + a\mu^2x_2^2 - 2a\mu^2x_1x_2 + 2a\mu x_1x_2 - 2a\mu x_2^2 \leq$$
$$a\mu x_1^2 - a\mu x_2^2 \quad \text{(A.10.5)}$$

$$\begin{aligned}
&\implies a\mu^2x_1^2 + a\mu^2x_2^2 - 2a\mu^2x_1x_2 + 2a\mu x_1x_2 - a\mu x_2^2 - a\mu x_1^2 \leq 0 \\
&\implies x_1^2\left(a\mu^2 - a\mu\right) + x_2^2\left(a\mu^2 - a\mu\right) - 2x_1x_2\left(a\mu^2 - a\mu\right) \leq 0 \\
&\implies \left(a\mu^2 - a\mu\right)(x_1 - x_2)^2 \leq 0 \\
&\implies a\mu(1-\mu)(x_1 - x_2)^2 \geq 0 \quad \text{(A.10.6)}
\end{aligned}$$

For the inequality in (A.10.6) to be true,

$$a \geq 0 \because \mu, 1 - \mu \geq 0, (x_1 - x_2)^2 \geq 0 \quad \text{(A.10.7)}$$

However, $a \neq 0$, since it is a quadratic function. Hence $a > 0$, for $f(x)$ to be convex.

A.11. The quadratic form

$$q(\mathbf{x}) \triangleq \mathbf{x}^\top\mathbf{A}\mathbf{x} + \mathbf{b}^\top\mathbf{x} + c \quad \text{(A.11.1)}$$

is convex iff $\mathbf{A}$ is positive semi-definite.

Solution: Consider two points $\mathbf{x_1}$ and $\mathbf{x_2}$, and a real constant $0 \leq \mu \leq 1$. Then,

$$\begin{aligned}
&\mu f(\mathbf{x_1}) + (1-\mu) f(\mathbf{x_2}) - f(\mu \mathbf{x_1} + (1-\mu)\mathbf{x_2}) \\
&= \left(\mu - \mu^2\right) \mathbf{x_1}^\top \mathbf{A} \mathbf{x_1} + \left(1 - \mu - (1-\mu)^2\right) \mathbf{x_2}^\top \mathbf{A} \mathbf{x_2} \\
&- 2\mu (1-\mu) \mathbf{x_1}^\top \mathbf{A} \mathbf{x_2} && \text{(A.11.2)} \\
&= \mu (1-\mu) \left(\mathbf{x_1}^\top \mathbf{A} \mathbf{x_1} - 2\mathbf{x_1}^\top \mathbf{A} \mathbf{x_2} + \mathbf{x_2}^\top \mathbf{A} \mathbf{x_2}\right) && \text{(A.11.3)} \\
&= \mu (1-\mu) (\mathbf{x_1} - \mathbf{x_2})^\top \mathbf{A} (\mathbf{x_1} - \mathbf{x_2}) && \text{(A.11.4)}
\end{aligned}$$

Since $\mathbf{x_1}$ and $\mathbf{x_2}$ are arbitrary, it follows from (A.11.4) that

$$\mu f(\mathbf{x_1}) + (1-\mu) f(\mathbf{x_2}) \geq f(\mu \mathbf{x_1} + (1-\mu)\mathbf{x_2}) \tag{A.11.5}$$

iff $\mathbf{A}$ is positive semi-definite, as required.

A.12. Let

$$\mathbf{M} \triangleq \begin{pmatrix} \mathbf{m_1} & \mathbf{m_2} \end{pmatrix} \tag{A.12.1}$$

$$\boldsymbol{\mu} \triangleq \begin{pmatrix} \mu_1 \\ -\mu_2 \end{pmatrix} \tag{A.12.2}$$

$$\tag{A.12.3}$$

The function

$$f(\boldsymbol{\mu}) \triangleq \left\|\mathbf{M}\boldsymbol{\mu} - \mathbf{x}\right\|^2 \tag{A.12.4}$$

is convex.

Solution: (A.12.4) can be expressed as

$$(\mathbf{M}\boldsymbol{\mu} - \mathbf{x})^\top (\mathbf{M}\boldsymbol{\mu} - \mathbf{x}) \tag{A.12.5}$$

Consider $\boldsymbol{\mu}_1$ and $\boldsymbol{\mu}_2$ and let $0 \leq \mu \leq 1$. Then,

$$f(\mu\boldsymbol{\mu}_1 + (1-\mu)\boldsymbol{\mu}_2) = \|\mathbf{M}(\mu\boldsymbol{\mu}_1 + (1-\mu)\boldsymbol{\mu}_2) - \mathbf{x}\| \tag{A.12.6}$$

$$= \|\mu(\mathbf{M}\boldsymbol{\mu}_1 - \mathbf{x}) + (1-\mu)(\mathbf{M}\boldsymbol{\mu}_2 - \mathbf{x})\| \tag{A.12.7}$$

$$\leq \mu\|\mathbf{M}\boldsymbol{\mu}_\mathbf{1} - \mathbf{x}\| + (1-\mu)\|\mathbf{M}\boldsymbol{\mu}_\mathbf{2} - \mathbf{x}\| \tag{A.12.8}$$

Where (A.12.8) follows from the triangle inequality.

A.13. Show that the quadratic programming problem

$$\min_{\mathbf{x}} g(\mathbf{x}) = \|\mathbf{x} - \mathbf{P}\|^2 \tag{A.13.1}$$

$$\text{s.t. } h(\mathbf{x}) = \mathbf{x}^\top\mathbf{V}\mathbf{x} + 2\mathbf{u}^\top\mathbf{x} + f = 0, \quad \mathbf{V} \succeq \mathbf{0} \tag{A.13.2}$$

is nonconvex.

Solution: Let $\mathbf{x_1}$ and $\mathbf{x_2}$ satisfy (A.13.2). Then,

$$h(\mathbf{x}_1) = \mathbf{x_1}^\top\mathbf{V}\mathbf{x_1} + 2\mathbf{u}^\top\mathbf{x_1} + f = 0 \tag{A.13.3}$$

$$h(\mathbf{x}_2) = \mathbf{x_2}^\top\mathbf{V}\mathbf{x_2} + 2\mathbf{u}^\top\mathbf{x_2} + f = 0 \tag{A.13.4}$$

Then, for any $0 \leq \mu \leq 1$, from (A.11.4),

$$\mu h(\mathbf{x_1}) + (1-\mu)h(\mathbf{x_2}) = h(\mu\mathbf{x_1} + (1-\mu)\mathbf{x_2}) + \mu(1-\mu)(\mathbf{x_1} - \mathbf{x_2})^\top \mathbf{A}(\mathbf{x_1} - \mathbf{x_2}) \tag{A.13.5}$$

From (A.13.3) and (A.13.4),

$$\mu h\left(\mathbf{x_1}\right)+(1-\mu)\, h\left(\mathbf{x_2}\right)=0 \tag{A.13.6}$$

which, upon substituting in (A.13.5) yields

$$h\left(\mu\mathbf{x_1}+(1-\mu)\,\mathbf{x_2}\right)+\mu\,(1-\mu)\left(\mathbf{x_1}-\mathbf{x_2}\right)^{\top}\mathbf{A}\left(\mathbf{x_1}-\mathbf{x_2}\right)=0 \tag{A.13.7}$$

$$\implies h\left(\mu\mathbf{x_1}+(1-\mu)\,\mathbf{x_2}\right)\leq 0 \tag{A.13.8}$$

since

$$\mu\,(1-\mu)\left(\mathbf{x_1}-\mathbf{x_2}\right)^{\top}\mathbf{A}\left(\mathbf{x_1}-\mathbf{x_2}\right)>0 \tag{A.13.9}$$

Hence, the optimization problem is nonconvex as the set of points on the curve do not form a convex set.

A.14. If **P** lies *outside* the given curve, show that the following relaxation makes the above problem convex.

$$\min_{\mathbf{x}} g\left(\mathbf{x}\right)=\left\|\mathbf{x}-\mathbf{P}\right\|^{2} \tag{A.14.1}$$

$$\text{s.t. } h\left(\mathbf{x}\right)=\mathbf{x}^{\top}\mathbf{V}\mathbf{x}+2\mathbf{u}^{\top}\mathbf{x}+f\leq 0, \quad \mathbf{V}\succeq\mathbf{0} \tag{A.14.2}$$

Solution: In this case,

$$\mu h\left(\mathbf{x_1}\right)+(1-\mu)\, h\left(\mathbf{x_2}\right)\leq 0 \tag{A.14.3}$$

and from (A.13.5),

$$h\left(\mu \mathbf{x_1} + (1-\mu)\,\mathbf{x_2}\right) + \mu\,(1-\mu)\left(\mathbf{x_1} - \mathbf{x_2}\right)^\top \mathbf{A}\left(\mathbf{x_1} - \mathbf{x_2}\right) \leq 0 \tag{A.14.4}$$

$$\implies h\left(\mu \mathbf{x_1} + (1-\mu)\,\mathbf{x_2}\right) \leq 0 \tag{A.14.5}$$

Hence, the optimization problem is convex.

Appendix B

Gradient Descent

B.1. Find a numerical solution for (A.4.3)

(A.4.3)

Solution: A numerical solution for (A.4.3) is obtained as

$$\lambda_{n+1} = \lambda_n - \mu f'(\lambda_n) \tag{B.1.1}$$

$$= \lambda_n - \mu(2a\lambda_n + b) \tag{B.1.2}$$

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.

B.2. Write a program to implement (B.1.2).

Solution: Download and execute

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

B.3. Find a closed form solution for λ_n in (B.1.2) using the one sided Z transform.

B.4. Find the condition for which (B.1.2) converges, i.e.

$$\lim_{n\to\infty} |\lambda_{n+1} - \lambda_n| = 0 \tag{B.4.1}$$

Appendix C

Optimization

C.1. Lagrange Multipliers

C.1.1 Find

$$\min_{\mathbf{x}} f(\mathbf{x}) = \left\| \mathbf{x} - \begin{pmatrix} 8 \\ 6 \end{pmatrix} \right\|^2 = r^2 \tag{C.1.1.1}$$

$$\text{s.t.} \quad g(\mathbf{x}) = \begin{pmatrix} 1 & 1 \end{pmatrix} \mathbf{x} - 9 = 0 \tag{C.1.1.2}$$

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. C.1.1.1

```
manual/codes/2.1.py
```

C.1.2 Show that

$$\min r = \frac{5}{\sqrt{2}} \tag{C.1.2.1}$$

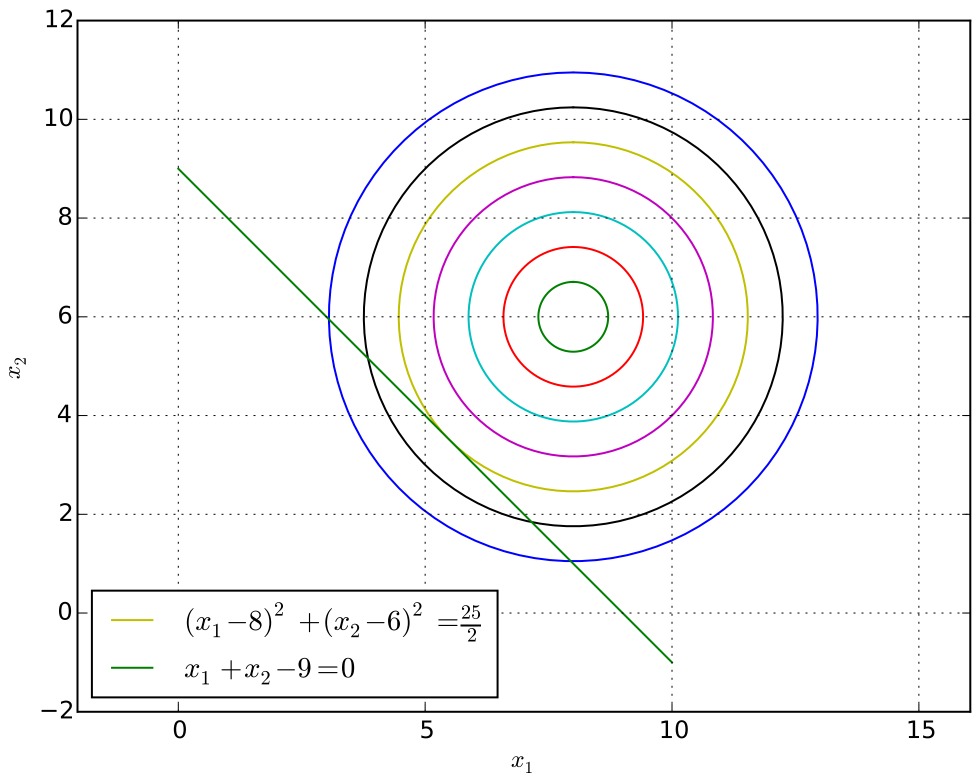


Figure C.1.1.1: Finding $\min_{\mathbf{x}} f(\mathbf{x})$

.

C.1.3 Show that

$$\nabla g(\mathbf{x}) = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \tag{C.1.3.1}$$

where

$$\nabla = \begin{pmatrix} \frac{\partial}{\partial x_1} \\ \frac{\partial}{\partial x_2} \end{pmatrix} \tag{C.1.3.2}$$

C.1.4 Show that

$$\nabla f(\mathbf{x}) = 2\left\{\mathbf{x} - \begin{pmatrix} 8 \\ 6 \end{pmatrix}\right\} \tag{C.1.4.1}$$

C.1.5 From Fig. C.1.1.1, show that

$$\nabla f(\mathbf{p}) = \lambda \nabla g(\mathbf{p}), \tag{C.1.5.1}$$

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

C.1.6 Use (C.1.5.1) and $\mathbf{g}(\mathbf{p}) = 0$ from (C.1.1.2) to obtain $\mathbf{p}$.

C.1.7 Define

$$L(\mathbf{x}, \lambda) = f(\mathbf{x}) - \lambda g(\mathbf{x}) \tag{C.1.7.1}$$

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

$$\nabla L(\mathbf{x}, \lambda) = 0. \tag{C.1.7.2}$$

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:

manual/codes/2.3.py

C.2. Inequality Constraints

C.2.1 Modify the code in problem C.1.1 to find a graphical solution for minimising

$$f(\mathbf{x}) \tag{C.2.1.1}$$

with constraint

$$g(\mathbf{x}) \geq 0 \tag{C.2.1.2}$$

Solution: This problem reduces to finding the radius of the smallest circle in the shaded area in Fig. C.2.1.1 . It is clear that this radius is 0.

```
manual/codes/2.4.py
```

C.2.2 Now use the method of Lagrange multipliers to solve problem C.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 C.2.1, which is different from the graphical solution. This means that the Lagrange multipliers method cannot be applied blindly.

C.2.3 Repeat problem C.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, $\lambda = 0$.

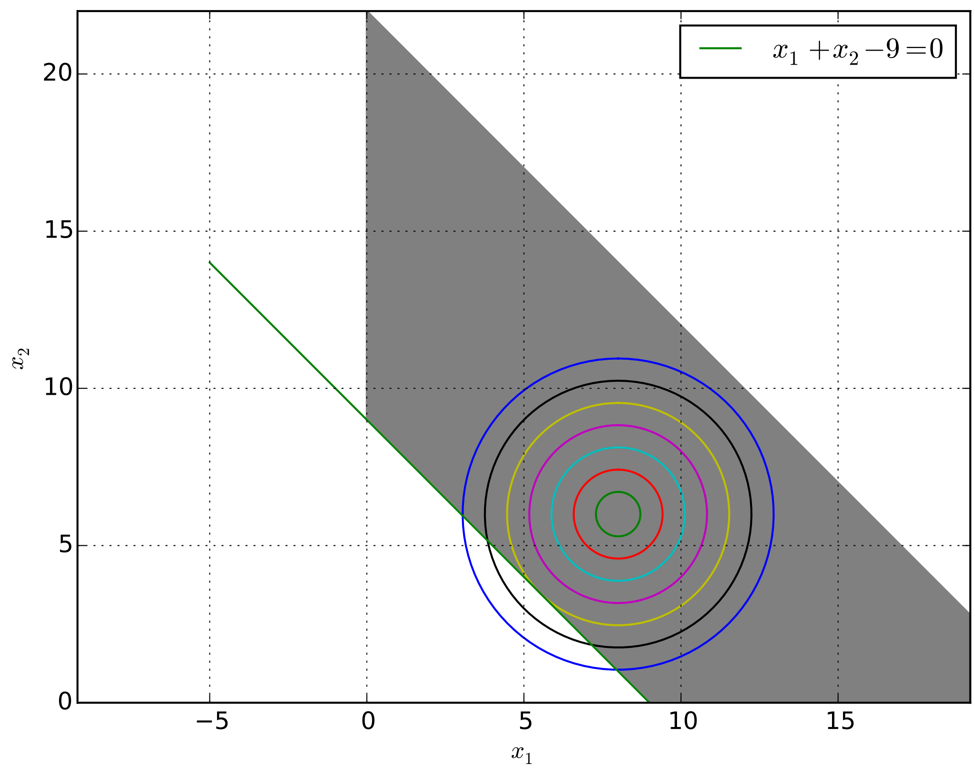


Figure C.2.1.1: Smallest circle in the shaded region is a point.

C.2.4 Find a graphical solution for minimising

$$f(\mathbf{x}) \tag{C.2.4.1}$$

with constraint

$$g(\mathbf{x}) \leq 0 \tag{C.2.4.2}$$

Summarize your observations.

Solution: In Fig. C.2.4.1, the shaded region represents the constraint. Thus, the solution is the same as the one in problem C.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 C.2.4.1 summarizes the conditions for this based on the observations so far.

manual/codes/2.7.py

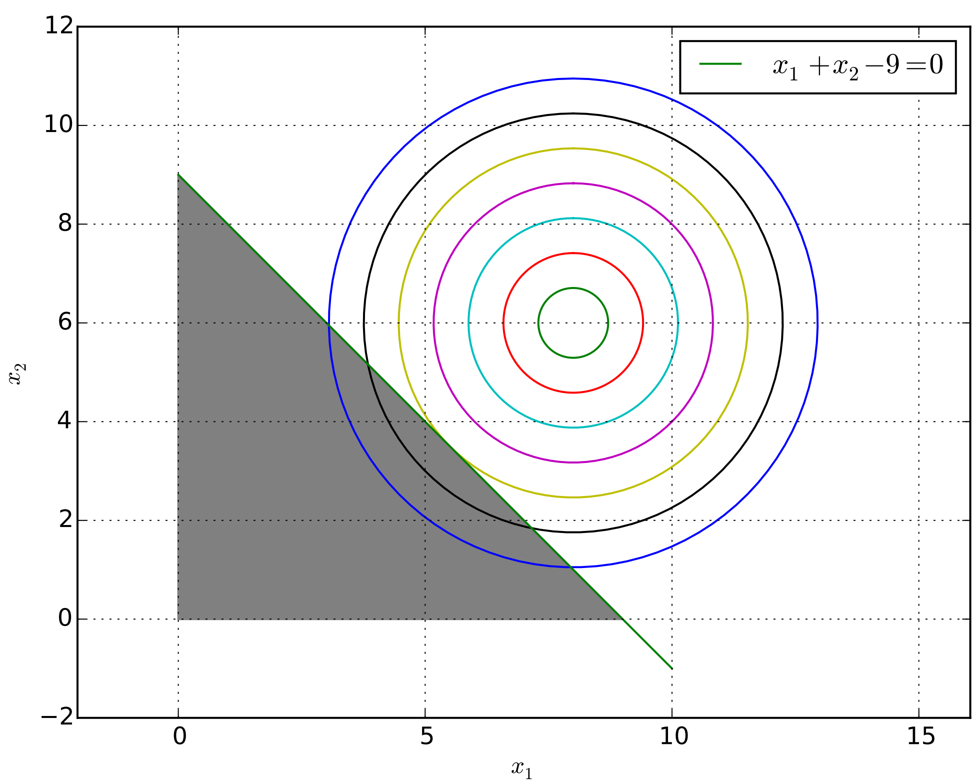


Figure C.2.4.1: Finding $\min_{\mathbf{x}} f(\mathbf{x})$.

Table C.2.4.1: Summary of conditions.

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

C.2.5 Find a graphical solution for

$$\min_{\mathbf{x}} f(\mathbf{x}) = \left\| \mathbf{x} - \begin{pmatrix} 8 \\ 6 \end{pmatrix} \right\|^2 \tag{C.2.5.1}$$

with constraint

$$g(\mathbf{x}) = \begin{pmatrix} 1 & 1 \end{pmatrix} \mathbf{x} - 18 = 0 \tag{C.2.5.2}$$

Solution:

```
manual/codes/2.8.py
```

C.2.6 Repeat problem C.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.

```
manual/codes/2.9.py
```

C.2.7 Solve

$$\min_{\mathbf{x}} f(\mathbf{x}) \tag{C.2.7.1}$$

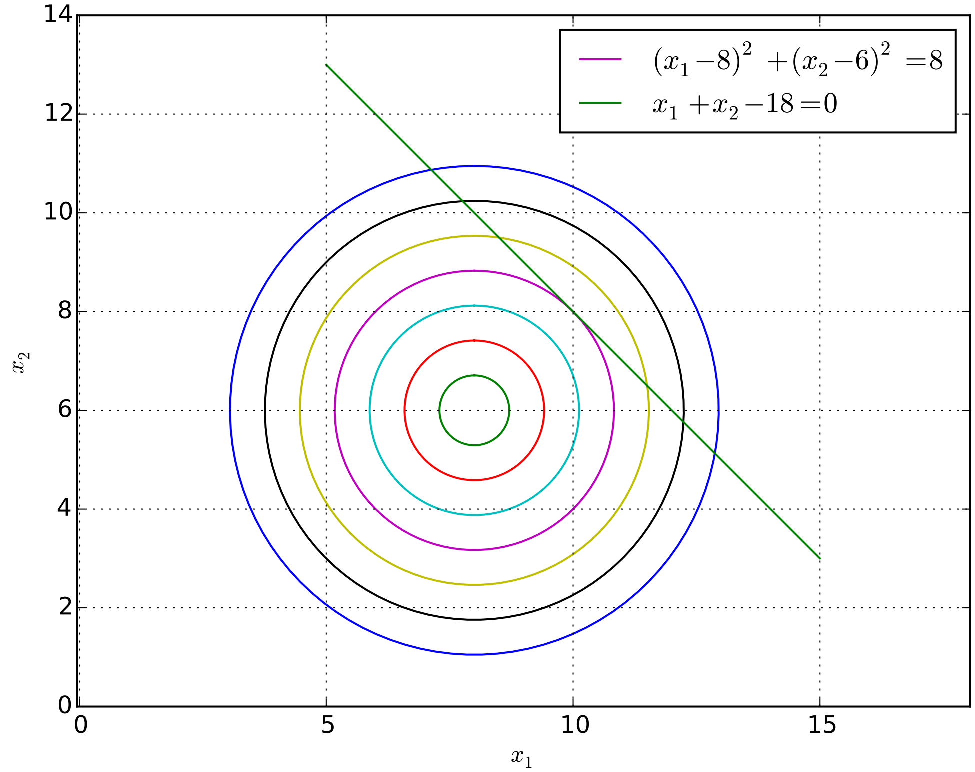


Figure C.2.5.1: Finding $\min_{\mathbf{x}} f(\mathbf{x})$.

with constraint

$$g(\mathbf{x}) \geq 0 \tag{C.2.7.2}$$

Solution: Since the unconstrained solution is outside the region $g(\mathbf{x}) \geq 0$, the solution is the same as the one in problem C.2.5.

C.2.8 Based on the problems so far, generalise the Lagrange multipliers method for

$$\min_{\mathbf{x}} f(\mathbf{x}), \quad g(\mathbf{x}) \geq 0 \tag{C.2.8.1}$$

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

$$\min_{\mathbf{x}} f(\mathbf{x}), \quad g(\mathbf{x}) \leq 0 \tag{C.2.8.2}$$

with the Lagrangian

$$L(\mathbf{x}, \lambda) = f(\mathbf{x}) + \lambda g(\mathbf{x}), \tag{C.2.8.3}$$

provided

$$\nabla L(\mathbf{x}, \lambda) = 0 \Rightarrow \lambda > 0 \tag{C.2.8.4}$$

else, $\lambda = 0$.

C.3. KKT Conditions

C.3.1 Solve

$$\min_{\mathbf{x}} f(\mathbf{x}) = \mathbf{x}^T \begin{pmatrix} 4 & 0 \\ 0 & 2 \end{pmatrix} \mathbf{x} \tag{C.3.1.1}$$

with constraints

$$g_1(\mathbf{x}) = \begin{pmatrix} 3 & 1 \end{pmatrix} \mathbf{x} - 8 = 0 \tag{C.3.1.2}$$

$$g_2(\mathbf{x}) = 15 - \begin{pmatrix} 2 & 4 \end{pmatrix} \mathbf{x} \geq 0 \tag{C.3.1.3}$$

Solution: Considering the Lagrangian

$$\nabla L(\mathbf{x}, \lambda, \mu) = 0 \tag{C.3.1.4}$$

resulting in the matrix equation

$$\Rightarrow \begin{pmatrix} 8 & 0 & 3 & 2 \\ 0 & 4 & 1 & 4 \\ 3 & 1 & 0 & 0 \\ 2 & 4 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \lambda \\ \mu \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 8 \\ 15 \end{pmatrix} \tag{C.3.1.5}$$

$$\Rightarrow \begin{pmatrix} x_1 \\ x_2 \\ \lambda \\ \mu \end{pmatrix} = \begin{pmatrix} 1.7 \\ 2.9 \\ -3.12 \\ -2.12 \end{pmatrix} \tag{C.3.1.6}$$

using the following python script. The (incorrect) graphical solution is available in Fig. C.3.1.1

```
manual/codes/2.12.py
```

Note that $\mu < 0$, contradicting the necessary condition in (C.2.8.4).

C.3.2 Obtain the correct solution to the previous problem by considering $\mu = 0$.

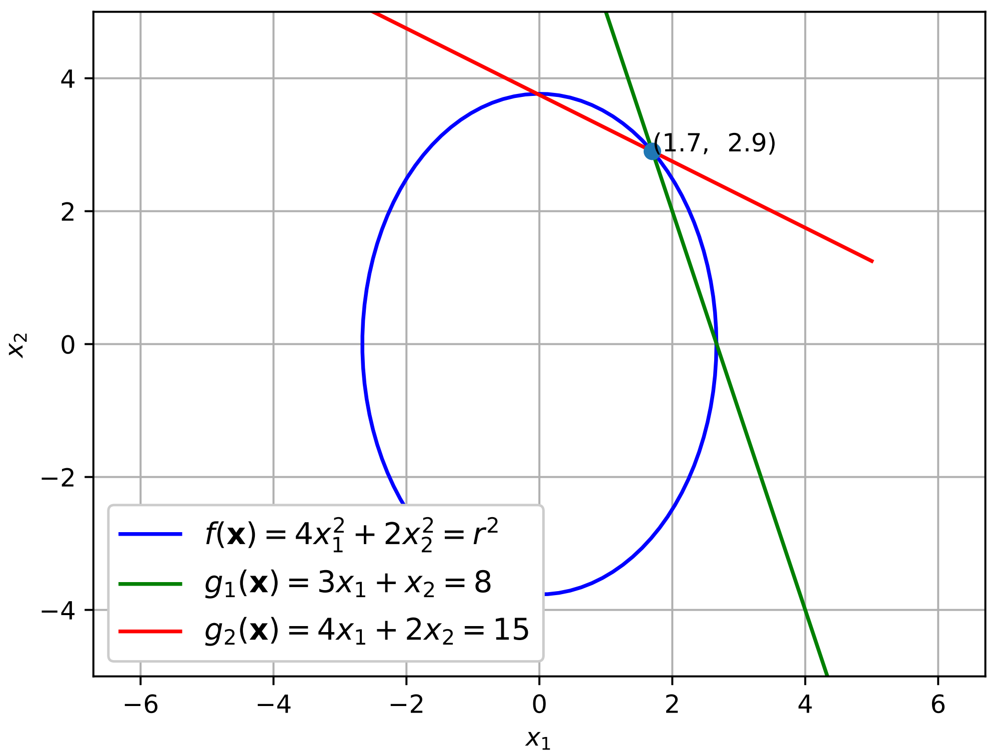


Figure C.3.1.1: Incorrect solution is at intersection of all curves $r = 5.33$

C.3.3 Solve

$$\min_{\mathbf{x}} f(\mathbf{x}) \tag{C.3.3.1}$$

with constraints

$$g_1(\mathbf{x}) = 0 \tag{C.3.3.2}$$

$$g_2(\mathbf{x}) \leq 0 \tag{C.3.3.3}$$

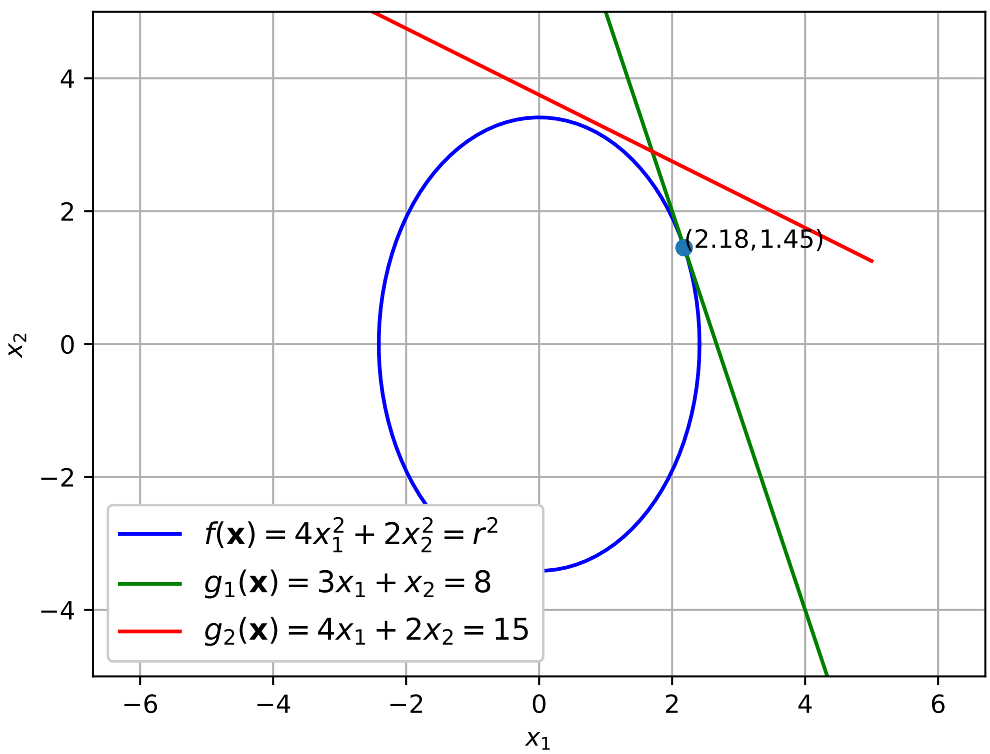


Figure C.3.2.1: Optimal solution is where $g_1(x)$ touches the curve $r = 4.82$

C.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 (C.3.1.1) using Lagrange Multipliers. These are called Karush-Kuhn-Tucker(KKT) conditions.

Solution: For a problem defined by

$$\mathbf{x}^* = \min_{\mathbf{x}} f(\mathbf{x}) \tag{C.3.4.1}$$

$$\text{subject to } h_i(\mathbf{x}) = 0, \forall i = 1, .., m \tag{C.3.4.2}$$

$$\text{subject to } g_i(\mathbf{x}) \leq 0, \forall i = 1, .., n \tag{C.3.4.3}$$

the optimal solution is obtained through

$$\mathbf{x}^* = \min_{\mathbf{x}} L(\mathbf{x}, \lambda, \mu) \tag{C.3.4.4}$$

$$= \min_{\mathbf{x}} f(\mathbf{x}) + \sum_{i=1}^{m} \lambda_i h_i(\mathbf{x}) + \sum_{i=1}^{n} \mu_i g_i(\mathbf{x}), \tag{C.3.4.5}$$

using the KKT conditions

$$\Rightarrow \nabla_{\mathbf{x}} f(\mathbf{x}) + \sum_{i=1}^{m} \nabla_{\mathbf{x}} \lambda_i h_i(\mathbf{x}) + \sum_{i=1}^{n} \mu_i \nabla_{\mathbf{x}} g_i(\mathbf{x}) = 0 \tag{C.3.4.6}$$

$$\text{subject to } \mu_i g_i(\mathbf{x}) = 0, \forall i = 1, .., n \tag{C.3.4.7}$$

$$\text{and } \mu_i \geq 0, \forall i = 1, .., n \tag{C.3.4.8}$$

C.3.5 Maxmimize

$$f(\mathbf{x}) = \sqrt{x_1 x_2} \tag{C.3.5.1}$$

with the constraints

$$x_1^2 + x_2^2 \leq 5 \tag{C.3.5.2}$$

$$x_1 \geq 0, x_2 \geq 0 \tag{C.3.5.3}$$

C.3.6 Solve

$$\min_{\mathbf{x}} \quad x_1 + x_2 \tag{C.3.6.1}$$

with the constraints

$$x_1^2 - x_1 + x_2^2 \leq 0 \tag{C.3.6.2}$$

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

Solution:

Graphical solution:

manual/codes/2.15.py

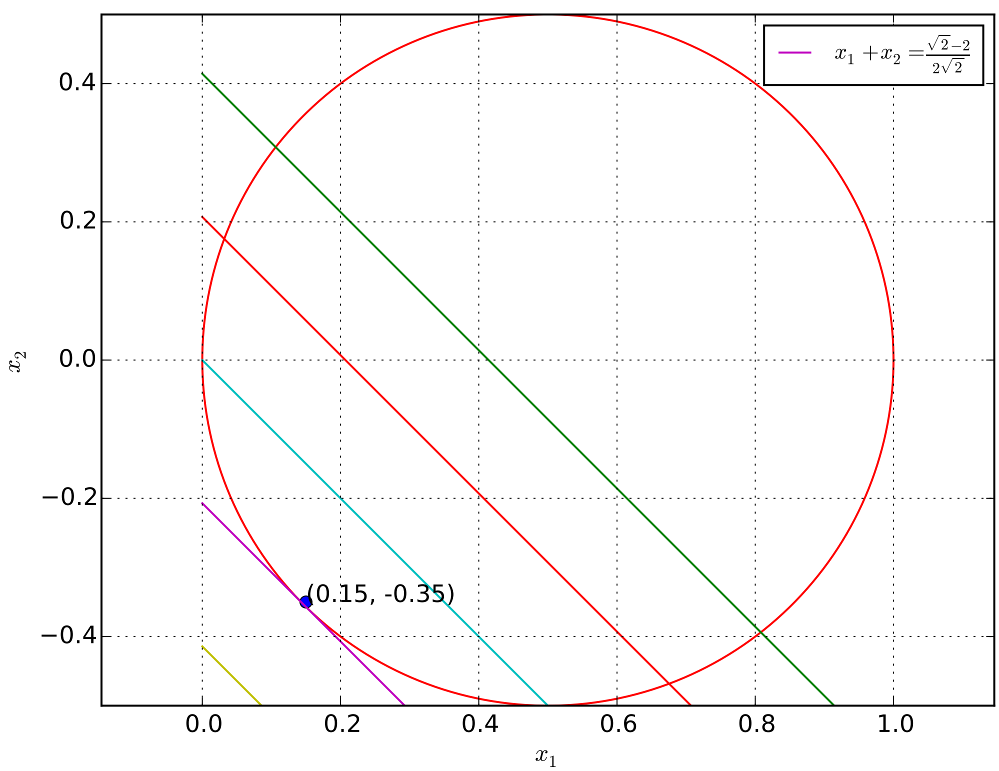


Figure C.3.6.1: Optimal solution is the lower tangent to the circle

Appendix D

Quadratic Programming

D.1. An apache helicopter of the enemy is flying along the curve given by

$$y = x^2 + 7 \tag{D.1.1}$$

A soldier, placed at

$$\mathbf{P} = \begin{pmatrix} 3 \\ 7 \end{pmatrix}. \tag{D.1.2}$$

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

$$\min_{\mathbf{x}} \|\mathbf{x} - \mathbf{P}\|^2 \tag{D.1.3}$$

$$\text{s.t. } \mathbf{x}^T\mathbf{V}\mathbf{x} + \mathbf{u}^T\mathbf{x} + d = 0 \tag{D.1.4}$$

where

$$\mathbf{V} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \tag{D.1.5}$$

$$\mathbf{u} = -\begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{D.1.6}$$

$$d = 7 \tag{D.1.7}$$

D.2. Show that the constraint in D.1.3 is nonconvex.

D.3. Show that the following <u>relaxation</u> makes (D.1.3) a convex optimization problem.

$$\min_{\mathbf{x}} \left(\mathbf{x} - \mathbf{P}\right)^T \left(\mathbf{x} - \mathbf{P}\right) \tag{D.3.1}$$

$$\text{s.t. } \mathbf{x}^T \mathbf{V} \mathbf{x} + \mathbf{u}^T \mathbf{x} \le 0 \tag{D.3.2}$$

D.4. Solve (D.3.1) using cvxpy.

Solution: The following code yields the minimum distance as 2.236 and the nearest point on the curve as

$$\mathbf{Q} = \begin{pmatrix} 1 \\ 8 \end{pmatrix} \tag{D.4.1}$$

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

D.5. Solve (D.3.1) using the method of Lagrange multipliers.

D.6. Graphically verify the solution to Problem D.1.

Solution: The following code plots Fig. D.6.1

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

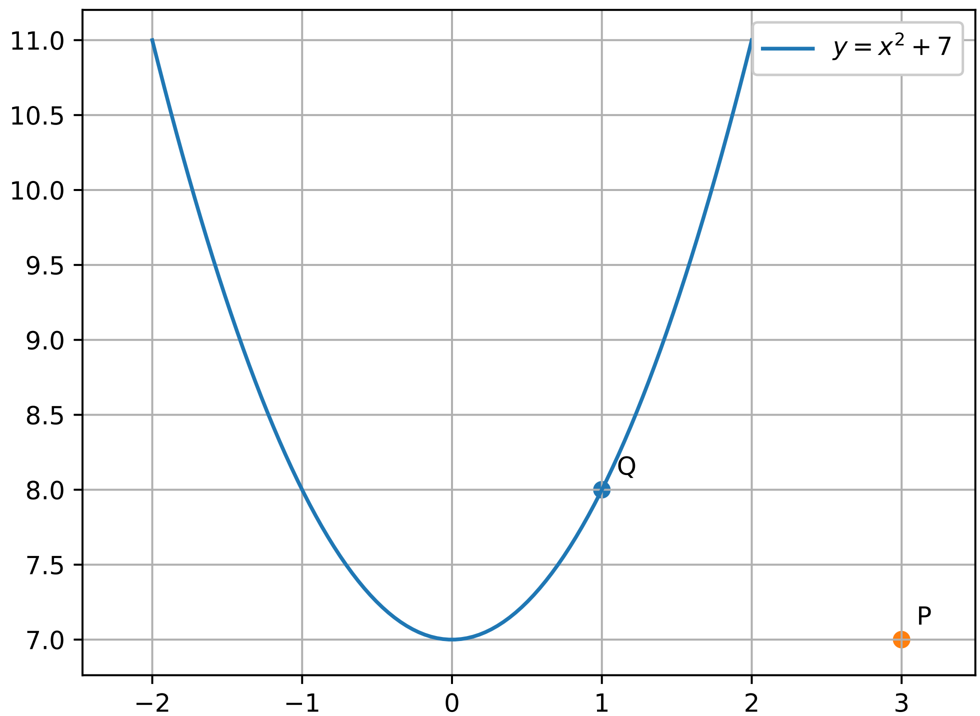


Figure D.6.1: **Q** is closest to **P**

.

D.7. Solve (D.3.1) using gradient descent.

Appendix E

Semi-Definite Programming

E.1 The problem

$$\min_{\mathbf{X}} x_{11} + x_{12} \tag{E.1.1}$$

with constraints

$$x_{11} + x_{22} = 1 \tag{E.1.2}$$

$$\mathbf{X} \succeq 0 \quad (\succeq \text{ means positive definite}) \tag{E.1.3}$$

where

$$\mathbf{X} = \begin{pmatrix} x_{11} & x_{12} \\ x_{12} & x_{22} \end{pmatrix} \tag{E.1.4}$$

is known as a semi-definite program. Find a numerical solution to this problem. Compare with the solution in problem C.3.6.

Solution: The cvxopt solver needs to be used in order to find a numerical solution.

For this, the given problem has to be reformulated as

$$\min_{\mathbf{x}} \begin{pmatrix} 1 & 1 & 0 \end{pmatrix} \begin{pmatrix} x_{11} \\ x_{12} \\ x_{22} \end{pmatrix} \quad \text{s.t} \tag{E.1.5}$$

$$\begin{pmatrix} 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_{11} \\ x_{12} \\ x_{22} \end{pmatrix} = 1 \tag{E.1.6}$$

$$x_{11} \begin{pmatrix} -1 & 0 \\ 0 & 0 \end{pmatrix} + x_{12} \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix} + x_{22} \begin{pmatrix} 0 & 0 \\ 0 & -1 \end{pmatrix} \preceq \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}. \tag{E.1.7}$$

The following script provides the solution to this problem.

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

E.2 Frame Problem E.1 in terms of matrices.

Solution: It is easy to verify that

$$x_{11} + x_{12} = \begin{pmatrix} 1 & 1 \end{pmatrix} \mathbf{X}^T \begin{pmatrix} 1 \\ 0 \end{pmatrix} \tag{E.2.1}$$

and

$$x_{11} + x_{22} = \begin{pmatrix} 1 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \mathbf{X} & \mathbf{0} \\ \mathbf{0} & \mathbf{X} \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix} \tag{E.2.2}$$

Thus, Problem E.1 can be expressed as

$$\begin{aligned} \min_{\mathbf{X}} \begin{pmatrix} 1 & 1 \end{pmatrix} \mathbf{X}^T \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad s.t \\ \begin{pmatrix} 1 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \mathbf{X} & \mathbf{0} \\ \mathbf{0} & \mathbf{X} \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix} = 1, \\ \mathbf{X} \succeq 0 \end{aligned} \tag{E.2.3}$$

E.3 Solve (E.2.3) using cvxpy.

Solution:

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

E.4 Minimize

$$-x_{11} - 2x_{12} - 5x_{22} \tag{E.4.1}$$

subject to

$$2x_{11} + 3x_{12} + x_{22} = 7 \tag{E.4.2}$$

$$x_{11} + x_{12} \geq 1 \tag{E.4.3}$$

$$x_{11}, x_{12}, x_{22} \geq 0 \tag{E.4.4}$$

$$\begin{pmatrix} x_{11} & x_{12} \\ x_{12} & x_{22} \end{pmatrix} \succeq 0 \tag{E.4.5}$$

using cvxpy.

E.5 Repeat the above exercise by converting the problem into a convex optimization problem in two variables and using graphical plots.

E.6 Solve the above problem using the KKT conditions. Comment.

Appendix F

Geometric Programming

F.1 A tank with rectangular base and rectangular sides, open at the top is to be constructed so that its depth is 2 m and volume is $8m^3$. If building of tank costs Rs 70 per sq metres for the base and Rs 45 per square metre for sides. What is the cost of least expensive tank?

Solution:

Let l,b and h be the length, width and height of a tank. The volume of tank is given by,

$$V = lbh \tag{F.1.1}$$

$$h = 2 \tag{F.1.2}$$

Cost of Building

$$R_b = 70/m^2 \tag{F.1.3}$$

$$R_s = 45/m^2 \tag{F.1.4}$$

The given problem can then be formulated as

$$S = \min_{l,b} R_b(lb) + R_s(4(l+b)) \tag{F.1.5}$$

$$\text{s.t} \quad lb = 4 \tag{F.1.6}$$

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

Appendix G

Line

G.1 The shortest distance between the lines whose vector equations are

$$L_1 : \mathbf{x} = \mathbf{x_1} + \lambda_1 \mathbf{m_1} \tag{G.1.1}$$

$$L_2 : \mathbf{x} = \mathbf{x_2} + \lambda_2 \mathbf{m_2} \tag{G.1.2}$$

is given by

$$\left\|\mathbf{A} - \mathbf{B}\right\| = \left\| \frac{\mathbf{m_1}^\top \mathbf{x} \mathbf{m_1}}{\left\|\mathbf{m_1}\right\|^2} - \mathbf{x} \right\| \tag{G.1.3}$$

Solution: Let $\mathbf{A}$ and $\mathbf{B}$ be points on lines L_1 and L_2 respectively such that AB is normal to both lines. Define

$$\mathbf{M} \triangleq \begin{pmatrix} \mathbf{m_1} & \mathbf{m_2} \end{pmatrix} \tag{G.1.4}$$

$$\boldsymbol{\lambda} \triangleq \begin{pmatrix} \lambda_1 \\ -\lambda_2 \end{pmatrix} \tag{G.1.5}$$

$$\mathbf{x} \triangleq \mathbf{x_2} - \mathbf{x_1} \tag{G.1.6}$$

Then, we have the following equations:

$$\mathbf{A} = \mathbf{x_1} + \lambda_1 \mathbf{m_1} \tag{G.1.7}$$

$$\mathbf{B} = \mathbf{x_2} + \lambda_2 \mathbf{m_2} \tag{G.1.8}$$

From (G.1.7) and (G.1.8), define the real-valued function f as

$$f(\boldsymbol{\lambda}) \triangleq \|\mathbf{A} - \mathbf{B}\| \tag{G.1.9}$$

$$= \|\mathbf{M}\boldsymbol{\lambda} - \mathbf{x}\| \tag{G.1.10}$$

From (A.12.4), since f is convex, differentiating

$$\|\mathbf{M}\boldsymbol{\lambda} - \mathbf{x}\|^2 = \boldsymbol{\lambda}^\top \mathbf{M}^\top \mathbf{M}\boldsymbol{\lambda} - 2\boldsymbol{\lambda}^\top \mathbf{M}^\top \mathbf{x} + \|x\|^2 \tag{G.1.11}$$

with respect to $\boldsymbol{\lambda}$ and equating to zero,

$$\mathbf{M}^\top (\mathbf{M}\boldsymbol{\lambda} - \mathbf{x}) = \mathbf{0} \tag{G.1.12}$$

yielding

$$\mathbf{M}^\top \mathbf{M}\boldsymbol{\lambda} = \mathbf{M}^\top \mathbf{x} \tag{G.1.13}$$

Appendix H

Manual