Chapter 5

A Trust-region Interior-point Technique to Solve Multiobjective Optimization Problems and its Application to a Tuberculosis Optimal Control Problem

5.1 Introduction

Pascoletti-Serafini [48] proposed a scalar optimization problem for determining (weakly) efficient solutions of a given MOP in terms of a given ordering cone. Compared with other existing scalarization techniques [46], this method has many advantages including handling nonconvexity [115]. In addition, some popular scalarization techniques such as weighted sum [31,116], ϵ-constraint [117], normal boundary intersection [40], ideal cone [47], etc. can be obtained by selecting suitable values of parameters and choosing an ordering cone. Recently, a modified Pascoletti-Serafini scalarization technique [118] has been established to generate the (weakly and properly) efficient solutions of the MOPs. Generally, the nature of converted scalar optimization problems by modified Pascoletti-Serafini scalarization technique are nonconvex. Therefore, these scalar optimization problems are hard to solve.

5.2 Motivation

The cone method requires an ideal point to decompose an MOP into single objective optimization subproblems. Calculating the ideal point of an MOP is itself a very difficult task. To avoid this extra computational work, we apply modified Pascoletti-Serafini scalarization technique to transform an MOP into scalar optimization subproblems.

To solve the transformed scalar optimization problems, an interior-point method that uses trust-region strategy is applied (see [119]). It enjoys the flexibility of switching between a line search method that computes steps by factoring the primal-dual equations and a trust region method that uses a conjugate gradient iteration. Steps computed by direct factorization are always tried first, but if they are deemed ineffective, a trust region iteration that guarantees progress toward stationarity is invoked. This method able to handle convex and nonconvex optimization problems.

5.3 Contributions

In this chapter, we introduce a trust-region interior-point technique to generate the Pareto optimal solutions for multiobjective optimization problems. The Pascoletti-Serafini scalarization technique is utilized to convert an MOP into a single-objective optimization subproblems. Thereafter, the subproblems are solved by a trust-region interior-point method. The performance of the proposed method is shown by applying it to some standard test problems. We also apply the proposed algorithm to solve an optimal control problem for a tuberculosis model [113].

The results of the proposed algorithm in this chapter are as follows.

(i) It is shown that the stationary points of a proposed merit function are also stationary points of the KKT conditions of the barrier problem.

(ii) It is also shown that the directions that are used to find the sequence of iterates of the proposed method are descent direction of the used merit function.

(iii) The implementaion of the proposed algorithm is tested on SCH, FON, ZDT1, ZDT2, ZDT3, ZDT4, DTLZ1 and DTLZ2 test problems. Also, the proposed algorithm is applied to solve an optimal control problem.

In this chapter, we consider the MOP of the following form:

$$\left.\begin{aligned} &\text{minimize} && F(x) = (f_1(x), f_2(x), \ldots, f_{\mathcal{P}}(x))^\top,\ \mathcal{P} \geq 2 \\ &\text{subject to} && h_j(x) \geq 0,\ j = 1, 2, \ldots, \mathcal{J}, \\ & && x \geq 0. \end{aligned}\right\} \tag{5.1}$$

5.4 Pascoletti-Serafini scalarization technique

In this section, a brief review of the modified Pascoletti-Serafini scalarization technique [118] is introduced. Let $a \in \mathbb{R}^{\mathcal{P}}_{\geqq}$ and $\mathrm{r} \in \mathbb{R}^{\mathcal{P}}_{\geqq} - \{0\}$. Then, the set of single objective nonlinear optimization subproblems, to solve (5.1), obtained by using the modified Pascoletti-Serafini scalarization technique is as follows:

$$\left.\begin{aligned} &\text{minimize} && t - \omega^\top \mathcal{G} \\ &\text{subject to} && a + t\mathrm{r} - F(x) - \mathcal{G} \geq 0 \\ & && h_j(x) \geq 0, \quad j = 1, 2, \ldots, \mathcal{P} \\ & && t \in \mathbb{R}, \quad x \geq 0, \quad \mathcal{G} \geq 0, \end{aligned}\right\} \tag{5.2}$$

where $\omega_i \geq 0, \quad i = 1, 2, \ldots, \mathcal{P}$ are weights and $\mathcal{G} = (\mathcal{G}_1, \mathcal{G}_2, \ldots, \mathcal{G}_{\mathcal{P}})^\top$.

To obtained the (weakly) efficient solutions of MOP (5.1), Pascoletti and Serafini suggested to solve the transformed scalar optimization problem (5.2) by changing the values of the parameters $a \in \mathbb{R}^{\mathcal{P}}_{\geqq}$, $\mathrm{r} \in \mathbb{R}^{\mathcal{P}}_{\geqq} - \{0\}$ and $\omega \geqq 0$.

We reformulate the problem (5.2) by introducing some notations. Let $\mathsf{c} = \big(0, 0, \ldots, 0, 1, -1\big)^\top \in \mathbb{R}^{n+2}$, $\mathsf{x} = (x_1, x_2, \ldots, x_n, x_{n+1}, x_{n+2})^\top$, $h(\mathsf{x}) = \big(h_1(\mathsf{x}), h_2(\mathsf{x}), \ldots, h_{\mathcal{J}}(\mathsf{x}), \mathsf{c}^\top\mathsf{x}\big)^\top$ and $t = x_{n+1} - x_{n+2} = \mathsf{c}^\top\mathsf{x}$. Then, the problem (5.2) becomes

$$\left.\begin{aligned} &\text{minimize} && \mathsf{c}^\top\mathsf{x} - \omega^\top\mathcal{G} \\ &\text{subject to} && a + \mathrm{r}\mathsf{c}^\top\mathsf{x} - F(\mathsf{x}) - \mathcal{G} \geq 0 \\ &&& h(\mathsf{x}) \geq 0 \\ &&& \mathsf{x} \geq 0, \quad \mathcal{G} \geq 0. \end{aligned}\right\} \tag{5.3}$$

We convert the inequality constraints of the problem (5.3) by introducing slack variables. Then, the problem (5.3) can be rewritten as follows:

$$\left.\begin{aligned} &\text{minimize} && \mathsf{c}^\top\mathsf{x} - \omega^\top\mathcal{G} \\ &\text{subject to} && \mathrm{G}(\mathsf{x}, \mathcal{G}) - s = 0 \\ &&& s \geq 0, \end{aligned}\right\} \tag{5.4}$$

where $\mathrm{G}(\mathsf{x}, \mathcal{G}) = \big(a + \mathrm{r}\mathsf{c}^\top\mathsf{x} - F(\mathsf{x}) - \mathcal{G}, h(\mathsf{x}), \mathsf{x}, \mathcal{G}\big)^\top$ and $s \in \mathbb{R}^{\mathcal{P}+(\mathcal{J}+1)+(n+2)+\mathcal{P}}$ is the slack variable.

In the forthcoming section, the procedure of solving the scalar optimization problem (5.4) by using an interior-point method is derived. Since the nature of the problem (5.4) is nonconvex, a variant of interior point method is used to solve problem (5.4).

5.5 Interior-point method with trust-region strategy

In this section, the scalar optimization problem (5.4) is solved with the help of an interior-point method with trust-region strategy. To make the problem (5.4) more straightforward, we introduce a new variable $\mathrm{z} = (\mathsf{x}, \mathcal{G})$. Now, we can rewrite problem (5.4) as follows:

$$\left.\begin{array}{ll} \text{minimize} & \mathrm{F}(\mathrm{z}) = \mathsf{c}^\top \mathsf{x} - \omega^\top \mathcal{G} \\ \text{subject to} & \mathrm{G}(\mathrm{z}) - s = 0 \\ & s \geq 0. \end{array}\right\} \tag{5.5}$$

For a given $\mu > 0$, a barrier problem associated to problem (5.5) is as follows:

$$\left.\begin{array}{ll} \text{minimize} & \mathcal{B}(\mathrm{z}, s; \mu) = \mathrm{F}(\mathrm{z}) - \mu \left(\displaystyle\sum_{i=1}^{n+2\mathcal{P}+\mathcal{J}+3} \log s_i \right) \\ \text{subject to} & \mathrm{G}(\mathrm{z}) - s = 0. \end{array}\right\} \tag{5.6}$$

The Lagrangian function associated with the barrier problem (5.6) is defined by

$$\mathcal{L}(\mathrm{z}, s, y; \mu) = \mathcal{B}(\mathrm{z}, s; \mu) + y^\top (\mathrm{G}(\mathrm{z}) - s), \tag{5.7}$$

where $y \in \mathbb{R}^{\mathcal{P}+(\mathcal{J}+1)+(n+2)+\mathcal{P}}$ is the Lagrange multiplier.

A sequential quadratic programming (SQP) method [33] with trust regions can solve the barrier problem (5.6) since it is an equality-constrained optimization problem. If SQP methods are directly applied to the barrier problem, they often produce very small step sizes that violate either or both of the positivity of the slack variables and the trust-region constraint. In order to address this issue, we design an SQP method that is tailored specifically to the structure of barrier problems.

We now apply the sequential quadratic programming method of [33] to the barrier problem (5.6). For a given barrier parameter $\mu > 0$, at an iterate (z, s), we first compute the Lagrange multiplier y (see Section 5.7) and then compute a step $d = (d_{\mathrm{z}}, d_s)^\top$ by solving the following quadratic programming problem:

$$
\left.\begin{aligned}
&\text{minimize} && \nabla \mathrm{F}^\top d_{\mathrm{z}} + \frac{1}{2} d_{\mathrm{z}}^\top \nabla^2_{\mathrm{zz}} \mathcal{L} d_{\mathrm{z}} - \mu e^\top S^{-1} d_s + \frac{1}{2} d_s^\top \Sigma d_s \\
&\text{subject to} && A(\mathrm{z})^\top d_{\mathrm{z}} - d_s + \mathrm{G}(\mathrm{z}) - s = \wp, \\
& && \|(d_{\mathrm{z}}, S^{-1} d_s)\| \leq \Delta, \\
& && d_s \geq -\tau s,
\end{aligned}\right\} \tag{5.8}
$$

where S is the diagonal matrix whose diagonal entries are components of the vector s; the scalar $\tau \in (0,1)$ is chosen close to 1 (for example, 0.99); $A(\mathrm{z})$ is the Jacobian of the function g and $\Sigma = \mu S^{-2}$; $\wp$ is the relaxation vector. The constraints $d_s \geq -\tau s$ maintain the positivity of d_s and the trust-region constraint $\|(d_{\mathrm{z}}, S^{-1} d_s)\| \leq \Delta$ guarantees that the problem (5.8) has a finite solution even when $\nabla^2_{\mathrm{zz}} \mathcal{L}(\mathrm{z}, s, y)$ is not positive definite. Therefore, the Hessian $\nabla^2_{\mathrm{zz}} \mathcal{L}(\mathrm{z}, s, y)$ need never be modified. In addition, the trust-region formulation ensures that adequate progress is made at every iteration. To justify the scaling S^{-1} that is used in the constraints of problem (5.8), we note that the shape of the trust region must take into account the requirement that the slacks not approach to zero prematurely. The scaling S^{-1} serves this purpose because it restricts those components i of the step vector d_s for which s_i is close to its lower bound zero.

In the next subsection, we provide a technique to compute the step d.

5.6 Step computation

The problem (5.8) is difficult to minimize exactly because of the presence of the nonlinear constraint $\|(d_{\mathrm{z}}, S^{-1} d_s)\| \leq \Delta$ and the bounds $d_s \geq -\tau s$. An important observation is that we can compute useful inexact solutions, at moderate cost. Since this approach scales up well with the number of variables and constraints, it provides a framework for developing practical interior-point methods for large-scale optimization.

The first step in the solution process is to make a change of variables that transforms the trust-region constraint $\|(d_{\mathrm{z}}, S^{-1}d_s)\| \leq \Delta$ into a ball. By defining

$$\tilde{d} = \begin{bmatrix} d_{\mathrm{z}} \\ \tilde{d}_s \end{bmatrix} = \begin{bmatrix} d_{\mathrm{z}} \\ S^{-1}d_s \end{bmatrix}, \tag{5.9}$$

we can now write the problem (5.8) as follows:

$$\left.\begin{aligned} &\text{minimize} \quad \nabla \mathrm{F}^\top d_{\mathrm{z}} + \frac{1}{2} d_{\mathrm{z}}^\top \nabla^2_{\mathrm{zz}} \mathcal{L} d_{\mathrm{z}} - \mu e^\top \tilde{d}_s + \frac{1}{2} \tilde{d}_s{}^\top S \Sigma S \tilde{d}_s \\ &\text{subject to} \quad A(\mathrm{z})^\top d_{\mathrm{z}} - S\tilde{d}_s + \mathrm{G}(\mathrm{z}) - s = \wp, \\ &\qquad\qquad\quad \|(d_{\mathrm{z}}, \tilde{d}_s)\| \leq \Delta, \\ &\qquad\qquad\quad \tilde{d}_s \geq -\tau e. \end{aligned}\right\} \tag{5.10}$$

The choice of the relaxation vector $\wp$ requires careful consideration as it impacts the efficiency of the method. Our goal is to choose $\wp$ as the smallest vector such that the constraints $A(\mathrm{z})^\top d_{\mathrm{z}} - S\tilde{d}_s + \mathrm{G}(\mathrm{z}) - s = \wp$, $\|(d_{\mathrm{z}}, \tilde{d}_s)\| \leq \Delta$ and $\tilde{d}_s \geq -\tau e$ are consistent for some reduced value of trust-region radius Δ. To compute the vector $\wp$, we solve the following problem in the variable $v = (v_{\mathrm{z}}, v_s)$:

$$\left.\begin{aligned} &\text{minimize} \quad \|A(w)^\top d_{\mathrm{z}} - S\tilde{d}_s + \mathrm{G}(\mathrm{z}) - s\|^2 \\ &\text{subject to} \quad \|(v_{\mathrm{z}}, v_s)\| \leq 0.8\Delta, \\ &\qquad\qquad\quad v_s \geq -(\tau/2)e. \end{aligned}\right\} \tag{5.11}$$

If we ignore the bound constraint $v_s \geq -(\tau/2)e$, this problem has the standard form of a trust-region problem [120], and we can compute an approximate solution by using the dogleg method [121]. If the solution violates the bounds $v_s \geq -(\tau/2)e$, we can reduce Δ so that these bounds are satisfied. After solving (5.11), we define the residuals $\wp$ as

$$\wp = A(\mathrm{z})^\top d_{\mathrm{z}} - S\tilde{d}_s + \mathrm{G}(\mathrm{z}) - s. \tag{5.12}$$

We are now ready to compute an approximate solution $\tilde{d}$ of the problem (5.10). By (5.12), the vector v is a particular solution of the linear constraints in problem (5.11). We can then solve the problem (5.10) by using the projected conjugate gradient (CG) [122]. We terminate the projected CG iteration by Steihaug's rules [122]: During the solution by CG we monitor the satisfaction of the trust-region constraint $\|(d_{\mathrm{z}}, \tilde{d}_s)\| \leq \Delta$ and stop if the boundary of this region is reached, if negative curvature is detected, or if an approximation solution is obtained. If the solution given by the projected CG iteration does not satisfy the bound constraints $\tilde{d}_s \geq -\tau e$, we backtrack so that they are satisfied. After the step $(d_{\mathrm{z}}, \tilde{d}_s)$ has been computed, we recover d from (5.9).

5.7 Choice of Lagrange multipliers and merit function

At every iteration, the choice of Lagrange multiplier y remains positive. Therefore, at an iterate (z, s), we choose y as

$$y = \min\left\{10^{-3}, \frac{\mu}{s_i}\right\}. \tag{5.13}$$

This choice of Lagrange multiplier y which is obtained in this manner will always be positive.

The role of the merit function is to decide whether a step is productive and should be accepted. Since interior-point methods can be seen as methods for solving the barrier problem (5.6), it is appropriate to define the merit function ϕ in terms of barrier functions. We may use, for example, an exact merit function of the form

$$\phi_\nu(\mathrm{z}, s) = \mathcal{B}(\mathrm{z}, s; \mu) + \nu\|\mathrm{G}(\mathrm{z}) - s\|, \tag{5.14}$$

where $\nu > 0$ is the penalty parameter.

The following theorem adresses the stationarity property of the merit function

(5.14).

Theorem 5.1 *Consider the barrier problem* (5.6). *For* $\mu > 0$, *let the point* $\bar{\lambda}_\mu = (\bar{\mathrm{z}}_\mu, \bar{s}_\mu)$ *be such that* $G(\bar{\mathrm{z}}_\mu) - \bar{s}_\mu = 0$. *Also, suppose the point* $\bar{\lambda}_\mu$ *is a stationary point of the merit function* $\phi_\nu(\mathrm{z}, s)$. *Then, the point* $\bar{\lambda}_\mu$ *is also a stationary point for the barrier function* $\mathcal{B}$.

Proof: The directional derivative of the merit function ϕ_ν at the point $\lambda_\mu = (\mathrm{z}_\mu, s_\mu)$ along the direction d is given by the following relation:

$$(\nabla\phi_\nu(\lambda_\mu))^\top d = (\nabla\mathcal{B}(\lambda_\mu))^\top d - \nu\|\mathrm{G}(\mathrm{z}_\mu) - s_\mu\|. \tag{5.15}$$

For any $\mu > 0$, $\|\mathrm{G}(\bar{\mathrm{z}}_\mu) - \bar{s}_\mu\| = 0$ and $\left(\nabla\phi_\nu(\bar{\lambda}_\mu)\right)^\top d = 0$. Then, from (5.15), we obtain $\left(\nabla\mathcal{B}(\bar{\lambda}_\mu)\right)^\top d = 0$. Hence, $\bar{\lambda}_\mu$ is the stationary point for the barrier function. □

In trust-region methods, the step d is valid if the following inequality holds:

$$\frac{\mathrm{pred}(d)}{\mathrm{ared}(d)} \leq \eta, \tag{5.16}$$

where $\eta > 0$, the *actual reduction* (ared) is defined as

$$\mathrm{ared}(d) = \phi_\nu(\mathrm{z}, s) - \phi_\nu(\mathrm{z} + d_\mathrm{z}, s + d_s) \tag{5.17}$$

and the *predicted reduction* (pred) is defined as

$$\mathrm{pred}(d) = q_\nu(0) - q_\nu(d), \tag{5.18}$$

where $q_\nu(d) = (\nabla\mathcal{B})^\top d + \frac{\sigma}{2}\left(d_\mathrm{z}^\top \nabla^2_{\mathrm{zz}}\mathcal{L} d_\mathrm{z} + \mu d_s^\top S^{-2} d_s\right) + \nu m(d)$, with

$$m(d) = \|A(\mathrm{z})d_\mathrm{z} - d_s + \mathrm{G}(\mathrm{z}) - s\| - \|\mathrm{G}(\mathrm{z}) - s\| \qquad \text{and} \quad \sigma = \begin{cases} 1 & \text{if } d_z^\top \nabla^2_{\mathrm{zz}}\mathcal{L} d_z > 0, \\ 0 & \text{otherwise.} \end{cases}$$

To determine an appropriate value of the penalty parameter ν, we require that ν be large enough that satisfies the following inequality for some $\rho \in (0, 1)$:

$$\mathrm{pred}(d) \geq \rho\nu \left(m(0) - m(d)\right). \tag{5.19}$$

The inequality (5.19) provides the minimum value of the penalty parameter ν. Therefore, the minimum value of the penalty parameter is determined as follows:

$$\nu_{\min} = \frac{\left(\nabla\mathcal{B}\right)^{\top} d + \frac{\sigma}{2}\left(d_{\mathrm{z}}^{\top}\nabla^2_{\mathrm{zz}}\mathcal{L}d_{\mathrm{z}} + \mu d_s^{\top} S^{-2} d_s\right)}{(1-\rho)\|\mathrm{G}(\mathrm{z}) - s\|}. \tag{5.20}$$

In the proposed method, we typically choose $\nu = 2\nu_{\min}$.

The following theorem ensures that the step d is descent direction for the merit function ϕ_ν.

Theorem 5.2 *Consider the barrier problem (5.6). Let the point $\Omega = (\mathrm{z}, s, y)$ be such that $s > 0$ and $y > 0$. Assume that $\mathcal{L}$, for (5.6), is positive definite. Then, for any $\mu > 0$ and $\nu > \nu_{\min}$ (see (5.20)), the step d that is computed by solving the problem (5.8) is a descent direction for the merit function ϕ_ν.*

Proof: Since $\nu > \nu_{\min}$, from (5.20), we get the following inequality:

$$\nu > \frac{\left(\nabla\mathcal{B}\right)^{\top} d + \frac{\sigma}{2}\left(d_{\mathrm{z}}^{\top}\nabla^2_{\mathrm{zz}}\mathcal{L}d_{\mathrm{z}} + \mu d_s^{\top} S^{-2} d_s\right)}{(1-\rho)\|\mathrm{G}(\mathrm{z}) - s\|}. \tag{5.21}$$

Note that $\nabla^2_{\mathrm{zz}}\mathcal{L}$ is positive definite and $\sigma \geq 0$. Therefore, $\frac{\sigma}{2}\left(d_{\mathrm{z}}^{\top}\nabla^2_{\mathrm{zz}}\mathcal{L}d_{\mathrm{z}} + \mu d_s^{\top} S^{-2} d_s\right) \geq 0$. Now, the inequality (5.21) can be written as

$$\left(\nabla\mathcal{B}\right)^{\top} d - v\|\mathrm{G}(\mathrm{z}) - s\| < -\rho\nu\|\mathrm{G}(\mathrm{z}) - s\|. \tag{5.22}$$

Also, the directional derivative of the merit function ϕ_ν along the direction d is given by

$$\left(\nabla\phi_\nu(\lambda_\mu)\right)^\top d = \left(\nabla\mathcal{B}\right)^\top d - v\|\mathrm{G}(\mathrm{z}) - s\|. \tag{5.23}$$

Since $\rho \in (0,1)$ and $\nu > 0$ then from (5.22) and (5.23), we obtain $\left(\nabla\phi_\nu(\lambda_\mu)\right)^\top d \leq -\rho\nu\|\mathrm{G}(\mathrm{z}) - s\| < 0$. Hence, the result follows. □

Algorithm 6 provides a detailed step-wise procedure to find the nondominated points of MOPs with the help of the process described above. We use the following error function for the stopping criteria:

$$E(\mathrm{z}, s, y) = \max\left\{\|c - (A(\mathrm{z}))^\top y\|, \|Sy - \mu e\|, \|\mathrm{G}(\mathrm{z}) - s\|\right\}. \tag{5.24}$$

5.8 Well-definedness of Algorithm 6

The well-definedness of Algorithm 6 depends on line numbers 8 and 9. The problem (5.11) is in the standard form of the trust-region problem if we neglect the constraint $\tilde{d}_s \geq -\tau e$ in the problem (5.11). By reducing the value of the trust-region Δ, dogleg method solves the problem (5.11) for v until the constraint $\tilde{d}_s \geq -\tau e$ is satisfied. Hence, at the kth iteration, Algorithm 6 will be able to compute v_k. The vector v is a particular solution of the constraints of the problem (5.10). CG method is applied to solve the equally-constrained quadratic problem (5.10) for $\tilde{d}$ in Algorithm 6. Thus, Algorithm 6 will be able to compute $\tilde{d}_k$ and hence the total step d_k (see (5.9)) at kth iteration.

As an example, we consider the following bi-objective optimization problem:

$$\left.\begin{array}{ll} \text{minimize} & \left((x_1-1)^2 + (x_1-x_2)^2, (x_1-x_2)^2 + (x_2-3)^2\right) \\ \text{subject to} & 0 \leq x_1 \leq 5, \quad 0 \leq x_2 \leq 5. \end{array}\right\} \tag{5.25}$$

The convergence of Algorithm 6 towards the Pareto points of the problem (5.25) is shown in Figure 5.1. To obtain a discrete approximation of Pareto front of the

Algorithm 6 Trust-region interior-point method for MOPs

Inputs:
(a) Given MOP (5.1)
(b) Provide the number of subproblems (N) to be solved

Initialization:
Set Pareto set $\mathcal{D} \leftarrow \emptyset$
Provide an initial point $w_0 = (\mathrm{z}_0, s_0, y_0) \in \mathbb{R}^{3n+2\mathcal{J}+5\mathcal{P}+8}$ with $\mathrm{z}_0 > 0, \quad s_0 > 0$
Provide the initial trust region radius $\Delta_0 > 0$ and the constants $0 < \sigma < 1$, $\eta > 0$, $0 < \rho < 1$
Choose an initial value for $\mu > 0$
Give the accuracy precision $\epsilon > 0$
Set $k \leftarrow 0$ (iteration number)

Main Part:
for $i = 1 : N$ **do**
  Choose randomly $a \in \mathbb{R}^{\mathcal{P}}, \mathrm{r} \in \mathbb{R}^{\mathcal{P}} \setminus \{0\}$ and $\omega \geq 0$
  while $E(\mathrm{z}_k, s_k, y_k) \geq \epsilon$ (see (5.24) for E) **do**
    while $E(\mathrm{z}_k, s_k, y_k; \mu_k) \geq \mu$ **do**
      Compute Lagrange multiplier y by solving the problem (5.13)
      Compute $v_k = (v_{\mathrm{z}}, v_s)$ by solving (5.11)
      Compute $\tilde{d}_k$ by applying the projected CG method to (5.10)
      Obtain the total step d_k from (5.9)
      Update ν_k to satisfy (5.19)
      Compute $\mathrm{pred}_k(d_k)$ by (5.18) and $\mathrm{ared}_k(d_k)$ by (5.17)
      if $\mathrm{ared}_k(d_k) \geq \eta \mathrm{pred}_k(d_k)$ **then**
        Set $\mathrm{z}_{k+1} \leftarrow \mathrm{z}_k + d_{\mathrm{z}}, \quad s_{k+1} \leftarrow s_k + d_s$
        Choose $\Delta_{k+1} = 1.2\Delta_k$
      else
        Set $\mathrm{z}_{k+1} = \mathrm{z}_k, \quad s_{k+1} = s_k$ and choose $\Delta_{k+1} = 0.8\Delta_k$
      end if
      $k \leftarrow k + 1$
    end while
    Set $\mu \leftarrow \sigma\mu$
  end while
  Calculate $f(\mathrm{z})$
  Update Pareto set $\mathcal{D} \leftarrow \mathcal{D} \bigcup \{f(\mathrm{z})\}$
end for
return The set $\mathcal{D}$ (a discrete approximation of the whole Pareto set)

problem (5.25), Algorithm 6 solves the Pascoletti-Serafini scalar optimization problem with different values of $a \in \mathbb{R}^p_{\geqq}$, $\mathrm{r} \in \mathbb{R}^p_{\geqq} - \{0\}$ and $\omega \geq 0$. We provide an initial point $w^{(0)} = (\mathrm{z}^{(0)}, s^{(0)}, y^{(0)})$, where $\mathrm{z}^{(0)} = (0.5, 2, 1, 1)$, $y^{(0)} = s^{(0)} = (1, 1, \ldots, 1) \in \mathbb{R}^{13}$, initial trust region radius $\Delta_0 = 0.3$ and initial value of $\mu = 0.98$. Also, we choose

$\sigma = 0.95$, $\eta = 0.8$, $\rho = 0.75$ and the precision $\epsilon = 10^{-6}$. Algorithm 6 starts by choosing $a = (0, 0)$, $\mathrm{r} = (0.986, 0.002)$ and $\omega = (1, 1)$. The Algorithm 6 then computes $E(\mathrm{z}_0, s_0, y_0) = 317.809$ and checks the termination condition $E(\mathrm{z}_0, s_0, y_0) > \epsilon$. As this condition is true, it computes the Lagrange multiplier $y_0 = 10^{-3}$ by (5.13), and further computes the direction d_0. Thereafter, it increases or decreases the radius of trust region according to the condition (5.16). Algorithm 6 solves the barrier problem (5.6) for decreasing value of μ. Hence, Algorithm 6 gradually moves towards the solution of the barrier problem (5.6). After finding one efficient point, Algorithm 6 changes the value of a, r and ω and then converges to another efficient point. In a similar way, Algorithm 6 generates a set of efficient points by changing the values of a, r and ω. We solve the problem (5.25) by taking 50 different values of the parameters a, r and ω. We have shown all the iterations in the objective space (see Figure 5.1). The blue point $(2.5, 3.25)$ in Figure 5.1 is the starting point in the objective space and violet points are the generated Pareto points corresponding to different values of a, r and ω. Note that the initial point remains unchanged throughout the entire process (see Figure 5.1).

In the next section, we apply Algorithm 6 to some standard test problems and solve a tuberculosis optimal control problem.

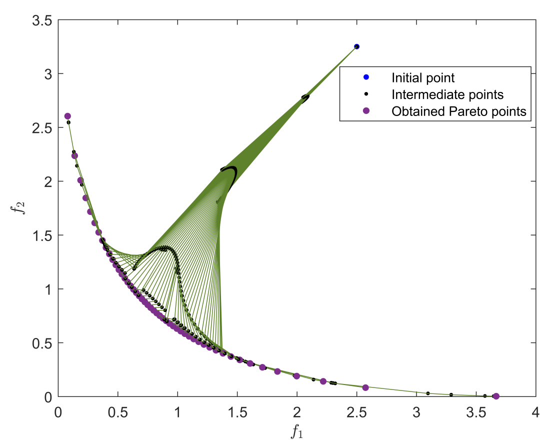

Figure 5.1: Obtained Pareto points of the example (5.25) by Algorithm 6

.

5.9 Numerical results and application

This subsection reports outcomes of the application of the proposed TR-IPM on several test problems (see Table 5.1) found in the literature. The test have been carried out on a PC with Intel Core i7-4770U 3.40 GHz CPU and 4GB RAM in MATLAB 2020a. The Pareto front of these problems are shown in red color and obtained Pareto points by Algorithm 6 shown in blue colour (see Figure 5.2). The graphs in Figure 5.2 show that the proposed algorithm can efficiently achieve convex, nonconvex, connected and disconnected Pareto fronts.

Table 5.1: Data for the test problems

Problem name	n	Nondominated set type	Variables' domain
SCH [97]	1	Convex	$[-1000, 1000]$
FON [12]	10	Nonconvex	$[-4, 4]$
ZDT1 [10]	10	Convex	$[0, 1]$
ZDT2 [10]	10	Nonconvex	$[0, 1]$
ZDT3 [10]	10	Nonconvex and disconnected	$[0, 1]$
ZDT4 [10]	10	Convex	$[0, 1]$
DTLZ1 [99]	10	Convex	$[0, 1]$
DTLZ2 [99]	10	Nonconvex	$[0, 1]$

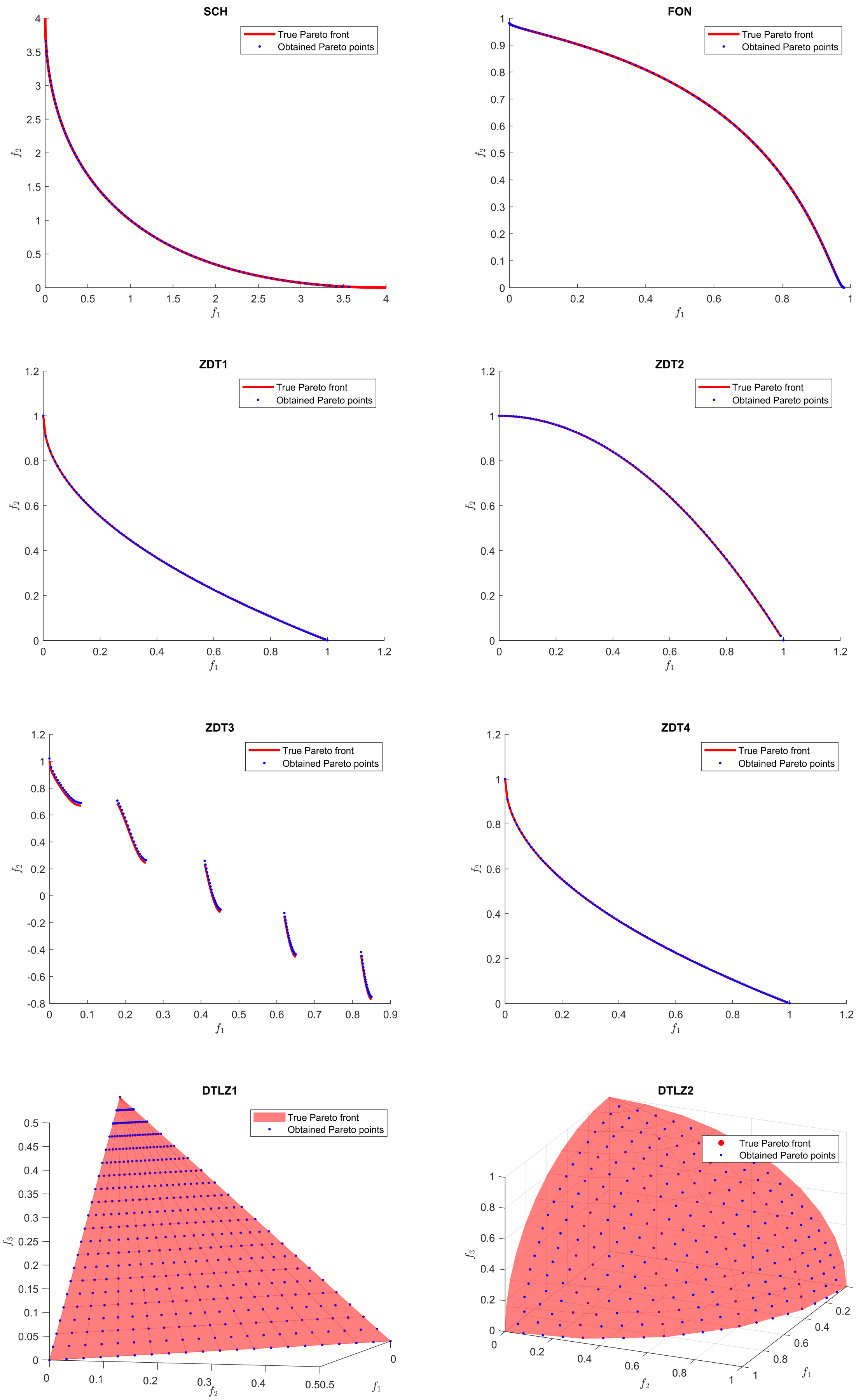

Figure 5.2: Obtained Pareto points vs true Pareto front.

5.10 Optimal control of a tuberculosis model

The current section provides an application of optimal control strategies in a tuberculosis (TB) model which is taken from [113]. The model is a system of ordinary differential equations which considers reinfection and post-exposure interventions. In the model, the population is divided into the following five categories:

- $\mathcal{S}$ represents the susceptible population;
- L_1 indicates the early latent population, i.e., a person who is recently infected for not more than two years but not infectious;
- I indicates the infected population which has active tuberculosis and is infectious;
- L_2 indicates the persistent latent population, i.e., the number of individuals who are infected and latent;
- R indicates the recovered population, i.e., the number of individuals who are previously infected and treated.

It is assumed that the total population $N = \mathcal{S} + L_1 + I + L_2 + R$ is constant. The control variable of the model system are as follows:

- $v_1(t)$ is a measurement representative that represents the measures taken to prevent treatment failure in active TB infected individuals I, such as following up with patients and helping them to take their prescribed medication and completing the treatment regimen.
- $v_2(t)$ represents the fraction of persistent latent individuals who have been identified and put under treatment.

According to Denysiuk et al. [113], the tuberculosis is modelled by the following system of nonlinear time-varying state equations with $v_1(t)$ and $v_2(t)$ as their control variable:

$$\left.\begin{aligned}
\dot{\mathcal{S}} &= \tilde{\mu}N - \frac{\beta}{N}I\mathcal{S} - \tilde{\mu}\mathcal{S},\\
\dot{L_1} &= \frac{\beta}{N}I(\mathcal{S} + \sigma_{L_2}L_2 + \sigma_R R) - (\sigma_{L_2} + \tau_1 + \tilde{\mu})L_1,\\
\dot{I} &= \phi_I \sigma_{L_2} L_1 + \omega L_2 + \omega_R R(t) - (\tau_0 + \epsilon_1 v_1(t) + \tilde{\mu})I,\\
\dot{L_2} &= (1-\phi_I)\delta_{L_1}L_1 - \sigma_{L_2}\frac{\beta}{N}IL_2 - (\omega + \epsilon_2 v_2(t) + \tau_2 + \tilde{\mu})L_2,\\
\dot{R} &= (\tau_0 + \epsilon v_1(t))I + \tau_1 L_1 + (\tau_2 + \epsilon_2 v_2(t))L_2 - \sigma_R\frac{\beta}{N}IR - (\omega_R + \tilde{\mu})R,
\end{aligned}\right\} \tag{5.26}$$

with the initial conditions

$$\mathcal{S}(0) = \frac{76}{120}N,\ L_1(0) = \frac{37}{120}N,\ I(0) = \frac{4}{120}N,\ L_2(0) = \frac{2}{120} \text{ and } N,\ R(0) = \frac{1}{120}N, \tag{5.27}$$

where the parameters in the system (5.26) are provided in the following table:

Symbols	Description	Value
β	Transmission coefficient	75, 100, 125
$\tilde{\mu}$	Death and birth rate	1/70 yr^{-1}
δ_{L_1}	Rate at which individuals leaves L_1	12 yr^{-1}
ϕ_I	Proportion of the individuals going to I	0.05
ω_L	Rate of endogenous reactivation for persistent latent infections	0.0002 yr^{-1}
ω_R	Rate of endogenous reactivation for treated individuals	0.00002 yr^{-1}
σ_{L_2}	Factor reducing the risk of infection as a result of acquired immunity to a previous infection for L_2	0.25
σ_R	Rate of exogenous reinfection of treated patients	0.25
τ_o	Rate of recovery under treatment of active TB	2 yr^{-1}
τ_1	Rate of recovery under treatment of latent individuals L_1	2 yr^{-1}
τ_2	Rate of recovery under treatment of latent individuals L_2	1 yr^{-1}
N	Total population	30,000
ϵ_1	Efficacy of treatment of active TB I	0.25
ϵ_2	Efficacy of treatment of active TB L_2	0.25
T	Total simulation duration	5 yr

Table 5.2: Description and values of the parameters in the model problem (5.32)

We obtain the values of parameters used in model system (5.26) from [113] and the references cited therein. Styblo [123] indicated that disease risk after infection is highest in the first five years and then declines exponentially. In the view of this, we consider the total simulation duration of $T = 5$ years.

5.11 Optimal control problem formulation

Our aim is to determine the optimal values v_1^* and v_2^* of the controls v_1 and v_2 such that (i) the corresponding state trajectories $\mathcal{S}^*$, L_1^* I^*, L_2^* and R^* are the solution of the system (5.26) and (5.27), in the time interval $[0, T]$ and (ii) minimize the following objective functional:

$$J(v_1, v_2) = \int_0^T \left[I(t) + L_2(t) + w_1 v_1^2(t) + w_2 v_2^2(t)\right] dt. \tag{5.28}$$

Here, the objective functional involves the number of active TB infectious individuals I, the number of persistent latent individuals L_2, and the implementation cost of the strategies associated with the controls v_1 and v_2. The controls are bounded by 0 and 1. When the controls vanish, no extra measures are implemented for the reduction of I and L_2; when the controls take the maximum value 1, the magnitude of the implemented measures associated to v_1 and v_2 take the values ϵ_1 and ϵ_2, respectively.

Consider the model system of ordinary differential equations (5.26). Suppose the set of permissible control functions is given by the following Lebesgue measurable set $\Psi = \{(v_1, v_2) : 0 \leq v_1(t), v_2(t) \leq 1,\ \forall t \in [0, T]\}$. Hence, the proposed problem is to minimize the objective functional (5.28), where the constants w_1 and w_2 give relative cost of the interventions associated to the control v_1 and v_2, respectively. The following problem

$$J(v_1^*, v_2^*) = \min_{(v_1, v_2) \in \Psi} J(v_1, v_2) \tag{5.29}$$

has two main aspects: (i) minimizing the number of active infected and latent individuals and (ii) reducing the implementation cost of control policies.

5.12 Multiobjective approach to the optimal control problem (5.29)

The available approaches that use the idea of optimal control theory have a significant disadvantage that it can only yield a single optimal solution for the problem (5.29). The determination of a single optimal requires decision-maker's preferences involving a pair of weight constants w_1 and w_2, for detail see [113], [123], and the references therein. However, the choice of parameters w_1 and w_2 is not straightforward. It requires some prior knowledge about the problem and decision maker's preferences, which may not always be available. Nevertheless, single optimal solution for the problem (5.29) does not provide all helpful insights for the optimal strategies and for the corresponding dynamics. Thus, many optimal alternatives remains unexplored by using the approaches based on optimal control theory. The current work proposes an approach that uses the idea of multiobjective optimization, which may provide the Pareto solutions instead of a single optimal solution. In the proposed approach, we decompose the cost functional (5.28) into two components f_1 and f_2, where

$$f_1(v_1, v_2) = \int_0^T \big(I(t) + L_2(t)\big)dt \text{ and} \tag{5.30}$$

$$f_2(v_1, v_2) = \int_0^T \big(v_1^2(t) + v_2^2(t)\big)dt. \tag{5.31}$$

Thus, a bi-objective formulation corresponding to (5.29) can be defined by the following:

$$\left.\begin{array}{ll}
\text{minimize} & (f_1, f_2) \\
\text{subject to} & \dot{\mathcal{S}} = \tilde{\mu}N - \frac{\beta}{N}I\mathcal{S} - \tilde{\mu}\mathcal{S}, \\
& \dot{L_1} = \frac{\beta}{N}I(\mathcal{S} + \sigma L_2 + \sigma_R R) - (\sigma + \tau_1 + \tilde{\mu})L_1, \\
& \dot{I} = \phi_I \sigma L_1 + \omega L_2 + \omega_R R(t) - (\tau_0 + \epsilon_1 v_1(t) + \tilde{\mu})I, \\
& \dot{L_2} = (1 - \phi_I)\delta_{L_1} L_1 - \sigma\frac{\beta}{N}IL_2 - (\omega + \epsilon_2 v_2(t) + \tau_2 + \tilde{\mu})L_2, \\
& \dot{R} = (\tau_0 + \epsilon v_1(t))I + \tau_1 L_1 + (\tau_2 + \epsilon_2 v_2(t))L_2 - \sigma_R\frac{\beta}{N}IR - (\omega_R + \tilde{\mu})R, \\
& \mathcal{S}(0) = \frac{76}{120}N,\ L_1(0) = \frac{37}{120}N,\ I(0) = \frac{4}{120}N,\ L_2(0) = \frac{2}{120}N,\ R(0) = \frac{1}{120}N, \\
& 0 \le v_1 \le 1, \\
& 0 \le v_2 \le 1.
\end{array}\right\} \tag{5.32}$$

Note that f_1 represents the number of active infected and latent individuals during the period $[0, T]$, and f_2 represents the cost associated to the implementation of control policies during the period $[0, T]$. In the above formulation, the constants w_1 and w_2 are ignored.

5.13 Algorithm

The following Algorithm 7 describes a complete procedure to obtain the Pareto set corresponding to the bi-objective optimal control problem (5.32). We first numerically integrate the model system (5.26) by using the fourth-order Runge-Kutta method on the interval $[0, T]$, where $T = 5$ years. We discretize the range of the controls v_1 and v_2 on 80 equally spaced time intervals. Then, the integrals in (5.30) and (5.31) are calculated by applying the trapezoidal rule. Further, using Algorithm 6 we find a discrete approximation of complete Pareto set of the problem (5.32).

Algorithm 7 Algorithm to obtain the Pareto set of the optimal control problem (5.32).

1: Compute the numerical integration of the system (5.26) by using the fourth-order Runge-Kutta method on the interval $[0, T]$.
2: Discretize the control v_1 and v_2 on a set of equally spaced time intervals.
3: Compute the integrals in (5.30) and (5.31) by applying the trapezoidal rule in order to obtain the discrete approximation of f_1 and f_2.
4: Compute a quadratic approximation of f_1 and f_2 with the help of curve fitting.
5: Apply Algorithm 6 for the bi-objective optimization problem (f_1, f_2) subject to $0 \leq v_1 \leq 1$, $0 \leq v_2 \leq 1$ to find the discrete approximation of complete Pareto set.

5.14 Numerical results

In the following section, we present and analyze the numerical results obtained by applying Algorithm 7 on the model system (5.26). The parameter values for the problem (5.26) are taken from Table 5.2. In the following part, we describe the obtained optimal solutions to the problem (5.32), taking into account the variations of the parameter β (transmission coefficient).

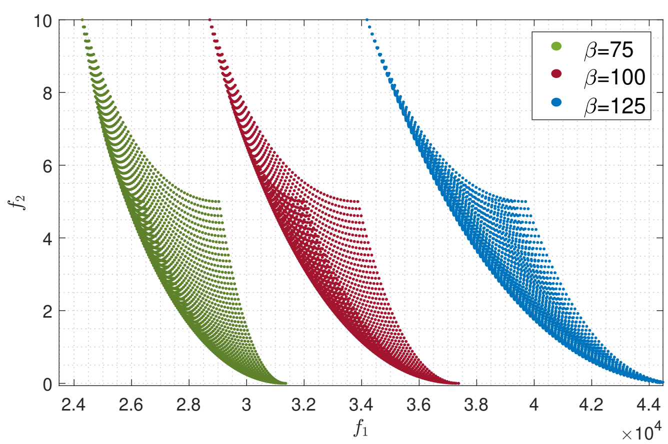

Figure 5.3: Feasible criterion region for different β's

Figure 5.3 plots the criterion feasible region of the problem (5.32) for $\beta = 75, 100$ and 125.

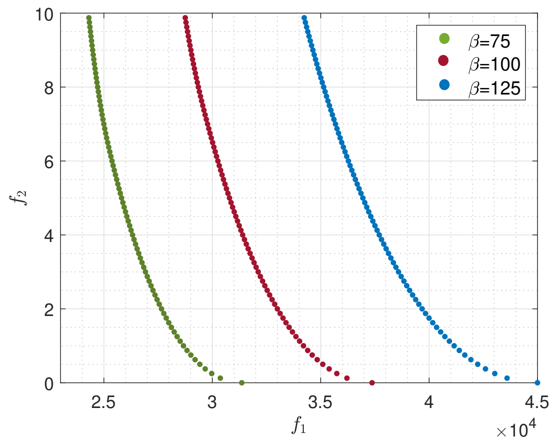

Figure 5.4: Trade-off curve for different values of β

Figure 5.4 plots the trade-off solution obtained for three different values of transmission coefficient β. From the figure, we observe that higher β values indicate higher level of infectious and persistent latent individuals is. Also, the difference between the worst and best case scenarios from medical perspective becomes large if β is increased.

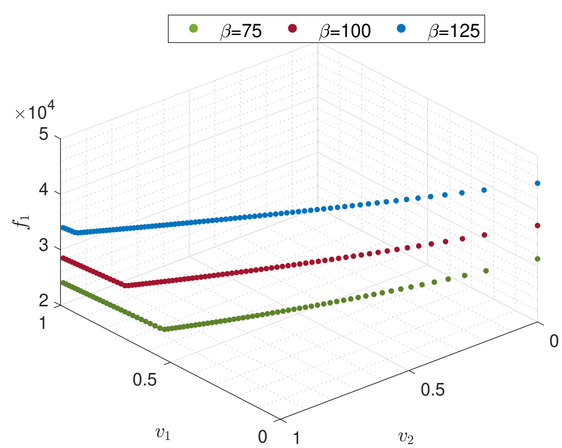

Figure 5.5: Representation of f_1 as a function of Pareto optimal strategies v_1 and v_2 for different values of β

In order to find the total number of active infected and latent individuals which are concentrated during the time span of 5 years we plot Figure 5.5. The figure depicts the values of f_1 as a function of Pareto optimal control strategy v_1^*, v_2^* for two different values of β. From Figure 5.5, we observe that the minimum and maximum values of f_1 increase as we increase the value of β.

5.15 Conclusion

In this chapter, multiobjective optimization problems have been solved with the help of Pascoletti-Serafini scalarization technique and interior point method. The Pascoletti-Serafini scalarization has been used to transform an MOP into a set of scalar optimization subproblems. Thereafter, these subproblems have been solved with the help of trust region interior point technique. The performance of Algorithm 6 has been tested on some standard test problems. The results in Section 5.9 has been shown the efficiency of Algorithm 6.

As an application, we have applied the proposed Algorithm to solve a model problem for tuberculosis from optimal control view point, using a multiobjective approach. The optimal control strategies are determined by simultaneously minimizing the number of individuals infected with tuberculosis and the cost of implementing prevention and treatment policies. Using the proposed approach, additional weight coefficients are not needed in order to formulate a single cost functional. The results of this study demonstrate how a comprehensive multiobjective approach can effectively identify the best control strategies within a mathematical model for tuberculosis. We found alternative viewpoints on implementing prevention and treatment policies using the obtained trade-off solutions. We analyze the optimal control strategies with varying β (transmission coefficient). The obtained results depict that as the transmission coefficient increases the proportion of active infectious and persistent latent individuals also increases. Furthermore, we have observed that the maximum and minimum values of f_1 increase with the increase of transmission coefficient β.

****