Mathematical Optimization in Radiotherapy Treatment Planning

Ehsan Salari

Department of Radiation Oncology
Massachusetts General Hospital and Harvard Medical School

HST S14
May 13, 2013

Outline

1. Introduction to Multi-criteria Optimization (MCO) in Radiotherapy Planning
2. Pareto-optimality and its Mathematical Characterization
3. MCO Solution Approaches for Radiotherapy Planning

Multi-criteria Optimization (MCO)

- Clinicians typically consider several treatment evaluation criteria when designing a radiotherapy plan
- These objectives can be conflicting
  - target coverage vs. organs-at-risk sparing
  - plan quality vs. delivery efficiency

Trade-off Between Treatment Evaluation Criteria

Treatment Planner

-----Original Message-----
From: Cornell, Mariel J.
Sent: Friday, October 26, 2007 8:21 AM
To: Hong, Theodore S.,M.D.
Subject: C***

Radiation Oncologist

Hi,
I played around with Ms C***'s plan all day yesterday, but even with relaxing the Rt Kidney restraints, the liver and stomach doses don't budge. If I push it, the PTV and CTV coverage really suffers. Are you willing to sacrifice their coverage or would you prefer to go with the plan you reviewed the other day? She's coming this afternoon for VSim.

Thanks,
Mariel

Trade-off Between Treatment Evaluation Criteria

- It may not be possible to satisfy all treatment objectives
  - clinicians may have to make compromise
  - it can be very time consuming
- The goal is to find a treatment plan that yields the desired trade-off between all evaluation criteria
- To achieve this goal, *multi-criteria optimization* (MCO) techniques are used
  - FMO problem [Küfer et al., 2003, Craft et al., 2006]
    - currently in clinical use
  - DAO problem [Salari and Unkelbach, 2013]
    - ongoing research

Trade-off Between Treatment Evaluation Criteria

- It may not be possible to satisfy all treatment objectives
  - clinicians may have to make compromise
  - it can be very time consuming
- The goal is to find a treatment plan that yields the desired trade-off between all evaluation criteria
- To achieve this goal, *multi-criteria optimization* (MCO) techniques are used
  - FMO problem [Küfer et al., 2003, Craft et al., 2006]
    - currently in clinical use
  - DAO problem [Salari and Unkelbach, 2013]
    - ongoing research

Trade-off Between Treatment Evaluation Criteria

- It may not be possible to satisfy all treatment objectives
  - clinicians may have to make compromise
  - it can be very time consuming
- The goal is to find a treatment plan that yields the desired trade-off between all evaluation criteria
- To achieve this goal, *multi-criteria optimization* (MCO) techniques are used
  - FMO problem [Küfer et al., 2003, Craft et al., 2006]
    - currently in clinical use
  - DAO problem [Salari and Unkelbach, 2013]
    - ongoing research

MCO Approach: Mathematical Formulation of FMO

- Mathematical formulation of the FMO problem

$$\min \left\{ G_1(\mathbf{d}), G_2(\mathbf{d}), \ldots, G_L(\mathbf{d}) \right\}$$

subject to

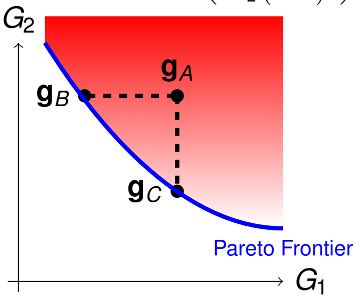
$$\begin{aligned} \mathbf{d} &= D^\top \mathbf{x} \\ H(\mathbf{d}) &\leq 0 \\ \mathbf{x} &\geq \mathbf{0} \end{aligned}$$

- Notation
  - $\mathbf{x}$: vector of beamlet intensities
  - $D = [D_{iv}]$: matrix of beamlet dose deposition coefficients
  - $\mathbf{d}$: vector of dose distribution
  - $\{G_\ell : \ell \in L\}$: collection of treatment evaluation criteria

Pareto Optimality

- In a multi-criteria optimization problem
  - notion of optimality is not well-defined
  - *Pareto-optimal* solutions are those for which we cannot improve any criteria unless some other criterion deteriorates

Treatment Plans $\mathbf{g} = \begin{pmatrix} G_1(D^\top \mathbf{x}) \\ G_2(D^\top \mathbf{x}) \end{pmatrix}$



Pareto Optimality: Mathematical Characterization

- Treatment plan $(\mathbf{x}^*, \mathbf{d}^*)$ *dominates* treatment plan $\left(\bar{\mathbf{x}}, \bar{\mathbf{d}}\right)$ if

$$
\begin{aligned}
G_\ell\left(\mathbf{d}^*\right) &\le G_\ell\left(\bar{\mathbf{d}}\right) & \forall \ell \in L \\
G_\ell\left(\mathbf{d}^*\right) &< G_\ell\left(\bar{\mathbf{d}}\right) & \exists \ell \in L
\end{aligned}
$$

- Dominated plans are not worth considering
- Treatment plan $(\mathbf{x}^*, \mathbf{d}^*)$ is Pareto optimal (Pareto efficient) if no other plan dominates it
- *Pareto frontier* is the collection of all Pareto-optimal plans

Characterizing Pareto-optimal Solutions

- There are several approaches to obtain Pareto-optimal solutions to MCO problems
  - see [Vira and Haimes, 2008]
- We discuss
  - weighted-sum method
  - ϵ-constraint method
- These techniques can generate all Pareto-optimal solutions to a convex problem

Weighted-sum Method

- MCO problem is transformed into a single-criterion problem using the *weighted-sum* objective function

$$\min \sum_{\ell \in L} w_\ell G_\ell(\mathbf{d})$$

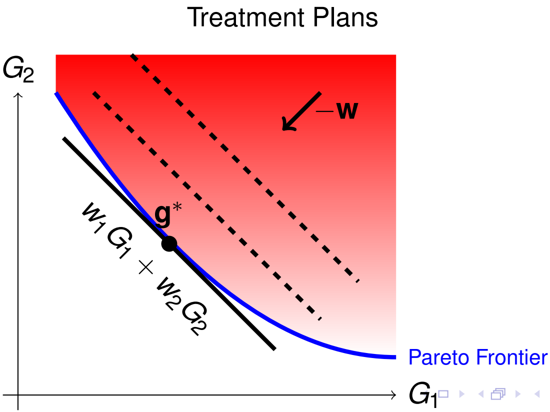
subject to

$$\begin{aligned} \mathbf{d} &= D^\top \mathbf{x} \\ H(\mathbf{d}) &\leq 0 \\ \mathbf{x} &\geq \mathbf{0} \end{aligned}$$

- Notation
  - **w**: vector of relative importance weights

Weighted-sum Method

- By solving the single-criterion problem for different *nonnegative* weight vectors **w** all Pareto-optimal solutions to a convex problem can be generated



ϵ-constraint Method

- MCO problem is transformed into a single-criterion problem by transforming all but one of the objectives, into constraints

$$\min G_{\hat{\ell}}(\mathbf{d})$$

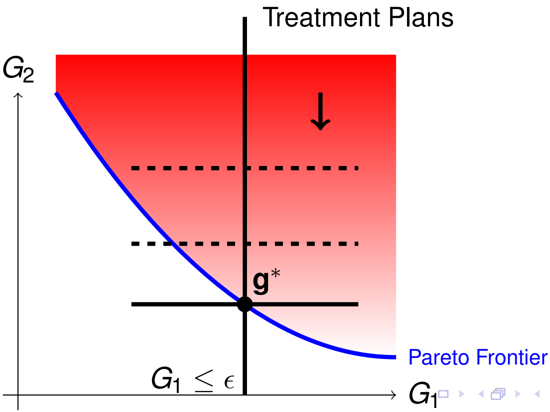
subject to

$$\begin{aligned} G_\ell(\mathbf{d}) &\leq \epsilon_\ell & \ell \in L \setminus \left\{\hat{\ell}\right\} \\ \mathbf{d} &= D^\top \mathbf{x} \\ H(\mathbf{d}) &\leq 0 \\ \mathbf{x} &\geq \mathbf{0} \end{aligned}$$

- Notation
  - ϵ: vector of right-hand-sides

ϵ-constraint Method

- By solving the single-criterion problem for different ϵ values all Pareto-optimal solutions to a convex problem can be generated



MCO Solution Methods

- There are three major MCO approaches to choose the desired Pareto-optimal solution?
  - ***a-priori* methods**: *preference* information exists prior to solving the MCO problem
  - ***interactive* methods**: it is an iterative process in which decision maker interacts with the solution method (e.g., via updating preference information)
  - ***a-posteriori* methods**: representative set of Pareto-optimal solutions (Pareto frontier) is generated and desired solution is chosen from the set

MCO Solution Methods

- There are three major MCO approaches to choose the desired Pareto-optimal solution?
  - ***a-priori* methods**: *preference* information exists prior to solving the MCO problem
  - ***interactive* methods**: it is an iterative process in which decision maker interacts with the solution method (e.g., via updating preference information)
  - ***a-posteriori* methods**: representative set of Pareto-optimal solutions (Pareto frontier) is generated and desired solution is chosen from the set

MCO Solution Methods

- There are three major MCO approaches to choose the desired Pareto-optimal solution?
  - ***a-priori* methods**: *preference* information exists prior to solving the MCO problem
  - ***interactive* methods**: it is an iterative process in which decision maker interacts with the solution method (e.g., via updating preference information)
  - ***a-posteriori* methods**: representative set of Pareto-optimal solutions (Pareto frontier) is generated and desired solution is chosen from the set

A-priori Method: Lexicographic Optimization

1. Initialize: Rank criteria $\ell \in L$ according to user's preference
2. Main step: At iteration $n \leq L$ solve

$$\min G_{\ell_{(n)}}(\mathbf{d})$$

subject to

$$G_{\ell}(\mathbf{d}) \leq G_{\ell}^{*} \qquad \ell = 1, \ldots, \ell_{(n-1)}$$
$$\mathbf{d} = D^{\top}\mathbf{x},\ H(\mathbf{d}) \leq 0,\ \mathbf{x} \geq \mathbf{0}$$

to obtain optimal objective value $G_{\ell_{(n)}}^{*}$

- Is final solution Pareto optimal?

Interactive Method: Scalarization Technique

1. Initialize: Assign weight vector $\mathbf{w}$ with equal weights to criteria (or according to user's preference if any)
2. Main step: Solve weighted-sum problem using $\mathbf{w}$
3. Termination condition: If desired trade-off is obtained, then stop; otherwise, modify $\mathbf{w}$ according to user's preference and go to Step 2

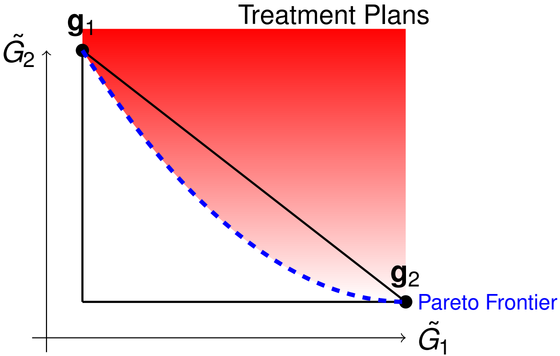
- Is final solution Pareto optimal? What is the disadvantage?

A-posteriori Method: Generating Pareto Frontier

- For majority of MCO problems we can only approximate Pareto frontier
  - by generating a collection of Pareto optimal points and use that to approximate Pareto frontier
- We discuss a *sandwich approximation* technique
  - see [Solanki et al., 1993, Vira and Haimes, 2008]
  - see [Craft et al., 2006] for application to radiotherapy planning

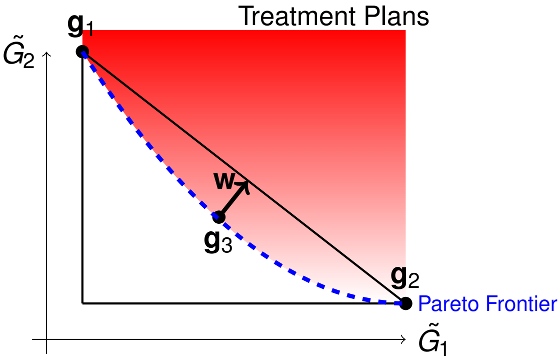
Bi-criteria Sandwich Approximation

- We first compute *anchor* points by minimizing each criterion individually (irrespective of others)
- We then normalize evluation criteria $\tilde{G}_\ell = \frac{G_\ell - G_\ell^{\min}}{G_\ell^{\max} - G_\ell^{\min}}$ $\ell \in L$



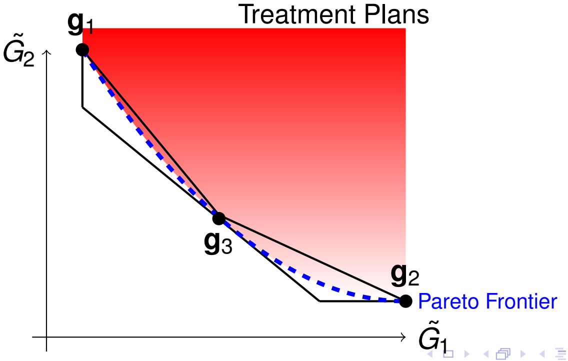
Bi-criteria Sandwich Approximation

- We then solve weighted-sum problem using $\mathbf{w} = (1, 1)$ to obtain new Pareto-optimal point $\mathbf{g}_3$



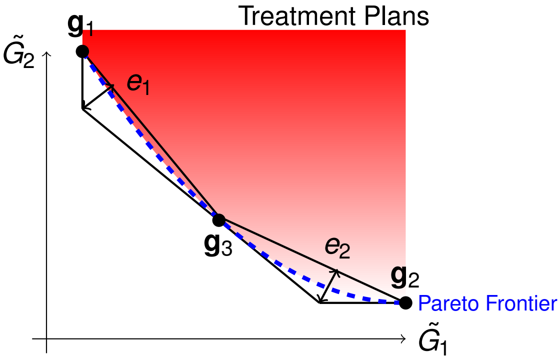
Bi-criteria Sandwich Approximation

- $\mathbf{x}_3$ is then used to update Pareto frontier approximation
- Supporting hyperplanes at $\{\mathbf{g}_1, \mathbf{g}_2, \mathbf{g}_3\}$ provide *outer approximation*
- Convexhull of $\{\mathbf{g}_1, \mathbf{g}_2, \mathbf{g}_3\}$ provides *inner approximation*



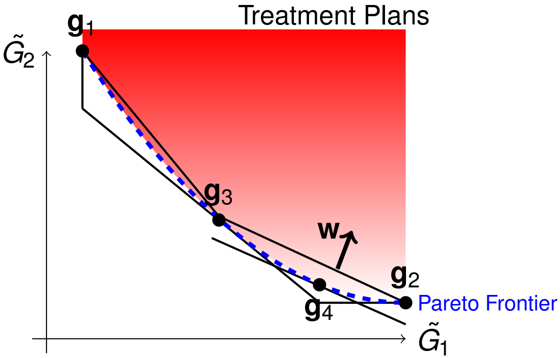
Bi-criteria Sandwich Approximation

- Maximum distance between inner and outer approximations is a measure of approximation *error*



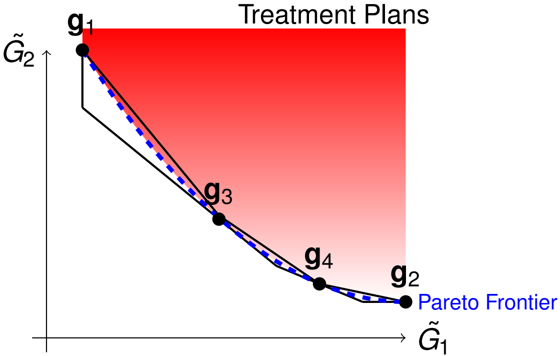
Bi-criteria Sandwich Approximation

- Using $\mathbf{w}$ = normal vector of inner approximation facet at maximum error, we solve weighted-sum problem to obtain new Pareto-optimal point $\mathbf{g}_4$



Bi-criteria Sandwich Approximation

- Using $\mathbf{g}_4$ we update inner and out approximations



Bi-criteria Sandwich Approximation

1. Initialize: Determine anchor and balanced solutions $X_{\mathcal{E}} = \{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3\}$ and normalize objectives G_1 and G_2
2. Maximum error: At iteration $n \geq 3$ determine Pareto segment with maximum approximation error

$$i^* = \underset{i}{\text{argmax}}\, e_i$$

3. Termination Condition: If $e_{i^*} < \epsilon$, stop; otherwise, go to Step 4
4. Main step: Let $\mathbf{w}_n$ be normal vector of inner approximation facet at maximum error and solve weighted-sum problem to obtain Pareto point $\mathbf{x}_n$
5. Update: Add $\mathbf{x}_n$ to $X_{\mathcal{E}}$ and update e_i $(i = 1, \ldots, n-1)$

Multi-criteria Sandwich Approximation

- We need to ensure that the weighted-sum problem has $\mathbf{w} \geq \mathbf{0}$
  - not an issue in bi-criteria case
- How to determine maximum distance and **w**?
  - inner approximation is formed by convex hull of Pareto points

$$Z_{in} = \left\{ \mathbf{z} = \sum_{i=1}^{n} \lambda_i \mathbf{g}_i : \sum_{i=1}^{n} \lambda_i = 1, \lambda_i \geq 0, \; i = 1, \ldots, n \right\}$$

  - outer approximation is formed by supporting hyperplanes

$$Z_{out} = \left\{ \mathbf{z} : \mathbf{w}_i^\top \mathbf{z} \geq \mathbf{w}_i^\top \mathbf{g}_i, \; i = 1, \ldots, n \right\}$$

Multi-criteria Sandwich Approximation

- Determining maximum distance between inner and outer approximations suggest by [Bokrantz and Forsgren, 2012]
  - Hausdorff distance between approximations

$$h\left(Z_{in}, Z_{out}\right) = \max_{\mathbf{z} \in Z_{out}} \min_{\mathbf{z}' \in Z_{in}} d\left(\mathbf{z}, \mathbf{z}'\right)$$

  - with one-sided distance function

$$d\left(\mathbf{z}, \mathbf{z}'\right) = \max_{\ell \in L} \left\{ \left(\mathbf{z}'_\ell - \mathbf{z}_\ell\right)_+ \right\}$$

Multi-criteria Sandwich Approximation

- Maximum distance can be obtained by solving the following linear bi-level programming problem

$$h(Z_{in}, Z_{out}) = \max_{\mathbf{z}} \left\{ \begin{array}{ll} \min_{\eta,\boldsymbol{\lambda}} & \eta \\ \text{s.t.} & \eta\,\mathbf{e} \geq \sum_{i=1}^{n} \lambda_i \mathbf{g}_i - \mathbf{z} \\ & \mathbf{e}^\top \boldsymbol{\lambda} = 1 \\ & \boldsymbol{\lambda} \geq \mathbf{0} \end{array} \right\}$$

subject to

$$\mathbf{w}_i^\top \mathbf{z} \geq \mathbf{w}_i^\top \mathbf{g}_i \qquad i = 1, \ldots, n$$

- Notation:
  - **e**: vector of ones

Multi-criteria Sandwich Approximation

- Optimal solution to this linear bi-level problem is among corner points of outer approximation
  - LP problem is solved for all corner points $\mathbf{z}_j$ for $j = 1, \ldots, n$

$$\min_{\eta,\boldsymbol{\lambda}} \eta$$

subject to (LP($\mathbf{z}_j$))

$$\eta\,\mathbf{e} \geq \sum_{i=1}^{n} \lambda_i \mathbf{g}_i - \mathbf{z}_j \qquad (\boldsymbol{\pi})$$

$$\mathbf{e}^\top \boldsymbol{\lambda} = 1$$

$$\boldsymbol{\lambda} \geq \mathbf{0}$$

- It can be shown that vector of Lagrangian multipliers $\boldsymbol{\pi}$ associated with inequality constraints is the normal vector

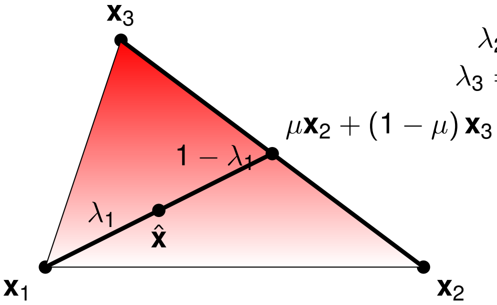
A-posteriori Method: Choosing Desired Trade-off

- Given Pareto frontier approximation and collection of Pareto-optimal solutions $X_{\mathcal{E}}$, how to choose solution with desired trade-off?
  - one can visualize Pareto frontier approximation for bi-criteria problems
  - can we devise an interactive exploration tool for multi-criteria problems ?
    - we discuss the approach implemented in RayStation®

A-posteriori Method: Convex Combination Plans

- One can search for a desired plan in the convex hull of Pareto-optimal solutions $X_{\mathcal{E}} = \{\mathbf{x}_n : n = 1, \ldots, N\}$

$$Conv\left(X_{\mathcal{E}}\right) = \left\{ \hat{\mathbf{x}} = \sum_{n=1}^{N} \lambda_n \mathbf{x}_n : \sum_{n=1}^{N} \lambda_n = 1, \boldsymbol{\lambda} \geq \mathbf{0} \right\}$$



$$\lambda_2 = \mu\left(1 - \lambda_1\right)$$
$$\lambda_3 = \left(1 - \mu\right)\left(1 - \lambda_1\right)$$

A-posteriori Method: Convex Combination Plans

- Mathematical properties of any solution in the convex hull
  - satisfies convex constraints enforced on MCO

$$\forall \hat{\mathbf{x}} \in Conv\left(X_{\mathcal{E}}\right) : \hat{\mathbf{d}} = D^{\top}\hat{\mathbf{x}},\ H\left(\hat{\mathbf{d}}\right) \leq 0,\ \hat{\mathbf{x}} \geq \mathbf{0}$$

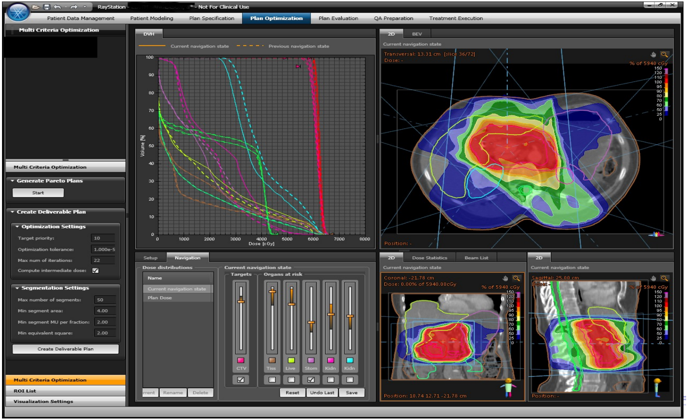
  - its dose distribution is convex combination of dose distributions of Pareto-optimal plans

$$\hat{\mathbf{x}} = \sum_{n=1}^{N} \hat{\lambda}_n \mathbf{x}_n \rightarrow \hat{\mathbf{d}} = \sum_{n=1}^{N} \hat{\lambda}_n \mathbf{d}_n$$

  - it is not necessarily Pareto optimal

A-posteriori Method: Exploring Convex Hull

- Sandwich approximation of Pareto frontier is obtained
- By interactively changing λ we can explore the convex hull of Pareto-optimal plans (so-called *database plans*)



Summary: MCO

- Clinicians consider a collection of conflicting criteria for treatment planning
- MCO is used to develop treatment planning approaches that allow for choosing desired trade-off
    - Pareto-optimality concept is used to quantify trade-off between these criteria
- There are three classes of MCO approaches:
    - a-priori, interactive, and a-posteriori methods
- A-posteriori method for radiotherapy planning consists of two stages:
    - stage I: Pareto frontier is approximated
    - stage II: plan with desired trade-off is chosen

References I

Bokrantz, R. and Forsgren, A. (2012).
An algorithm for approximating convex pareto surfaces based on dual techniques.
INFORMS Journal on Computing.

Craft, D. L., Halabi, T. F., Shih, H. A., and Bortfeld, T. R. (2006).
Approximating convex pareto surfaces in multiobjective radiotherapy planning.
Medical Physics, 33:3399–3407.

Küfer, K., Scherrer, A., Monz, M., Alonso, F., Trinkaus, H., Bortfeld, T., and Thieke, C. (2003).
Intensity-modulated radiotherapy–a large scale multi-criteria programming problem.
OR Spectrum, 25(2):223–249.

References II

Salari, E. and Unkelbach, J. (2013).
A column-generation-based method for multi-criteria direct aperture optimization.
Physics in Medicine and Biology, 58(3):621–639.

Solanki, R. S., Appino, P. A., and Cohon, J. L. (1993).
Approximating the noninferior set in multiobjective linear programming problems.
European Journal of Operational Research, 68(3):356–373.

Vira, C. and Haimes, Y. Y. (2008).
Multiobjective decision making: theory and methodology.
North-Holland.