

Optimization Modeling and Computational Issues in Radiation Therapy

(lecture developed in collaboration with Peng Sun)

February 5, 2002

Outline

1. Radiation Therapy
2. Linear Optimization Models
3. Computation
4. Nonlinear and Mixed-Integer Models
5. Looking Ahead to the Course

- This year, 1,200,000 Americans will be diagnosed with cancer
- 600,000+ patients will receive radiation therapy
 - beam(s) of radiation delivered to the body in order to kill cancer cells
- Sadly, only 67% of “curable” patients will be cured

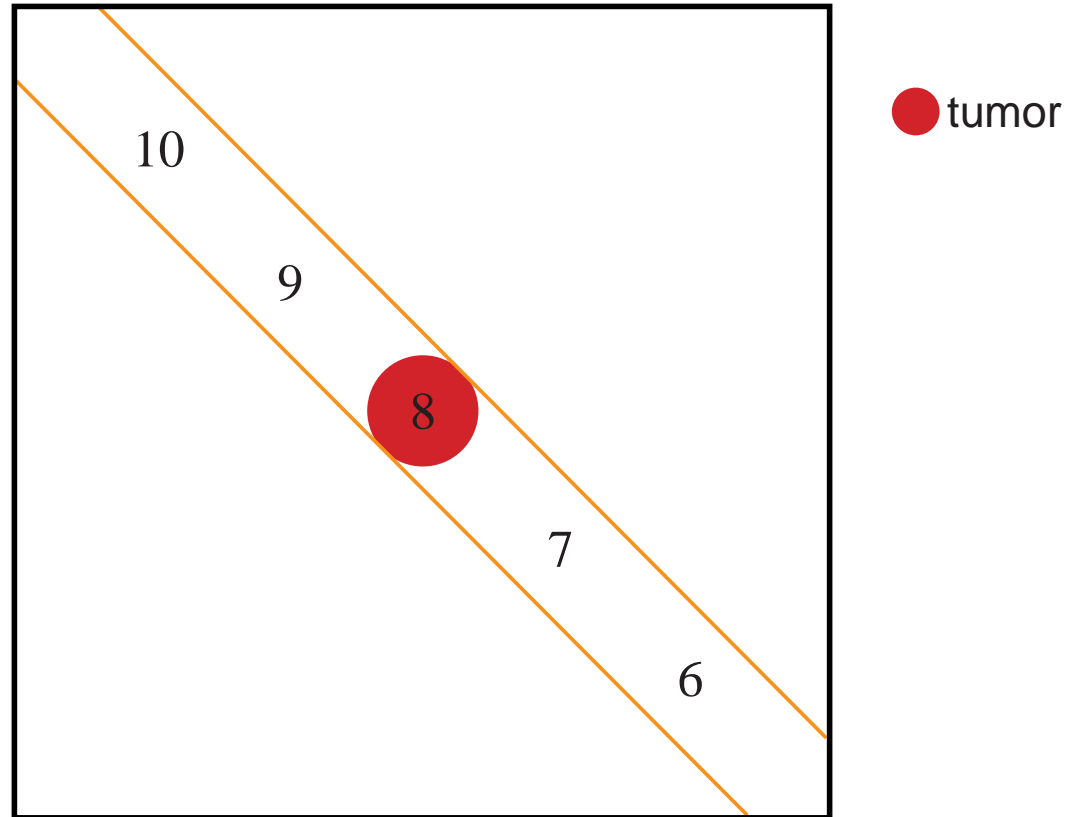
- High doses of radiation (energy/unit mass) can kill cells and/or prevent them from growing and dividing
 - true for cancer cells *and* normal cells
- Radiation is attractive because the repair mechanisms for cancer cells is less efficient than for normal cells

- Recent advances in radiation therapy now make it possible to:
 - map the cancerous region in greater detail
 - aim a larger number of different “beamlets” with greater specificity
- Spawned the new field of *tomotherapy*
- “Optimizing the Delivery of Radiation Therapy to Cancer Patients,” by Shepard, Ferris, Olivera, and Mackie, *SIAM Review*, Vol. 41, pp. 721–744, 1999.

Radiation Therapy

Overview

Conventional Radiotherapy...

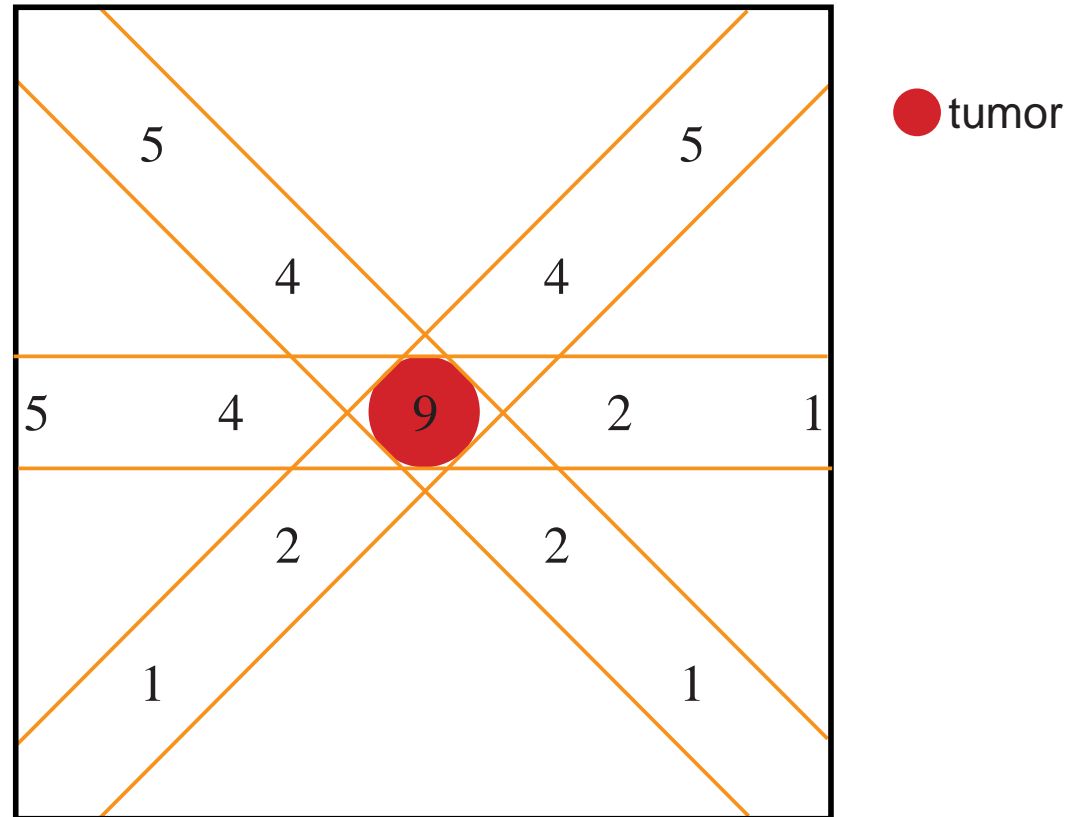


Relative Intensity of Dose Delivered

Radiation Therapy

Overview

...Conventional Radiotherapy...



Relative Intensity of Dose Delivered

Radiation Therapy

Overview

...Conventional Radiotherapy...

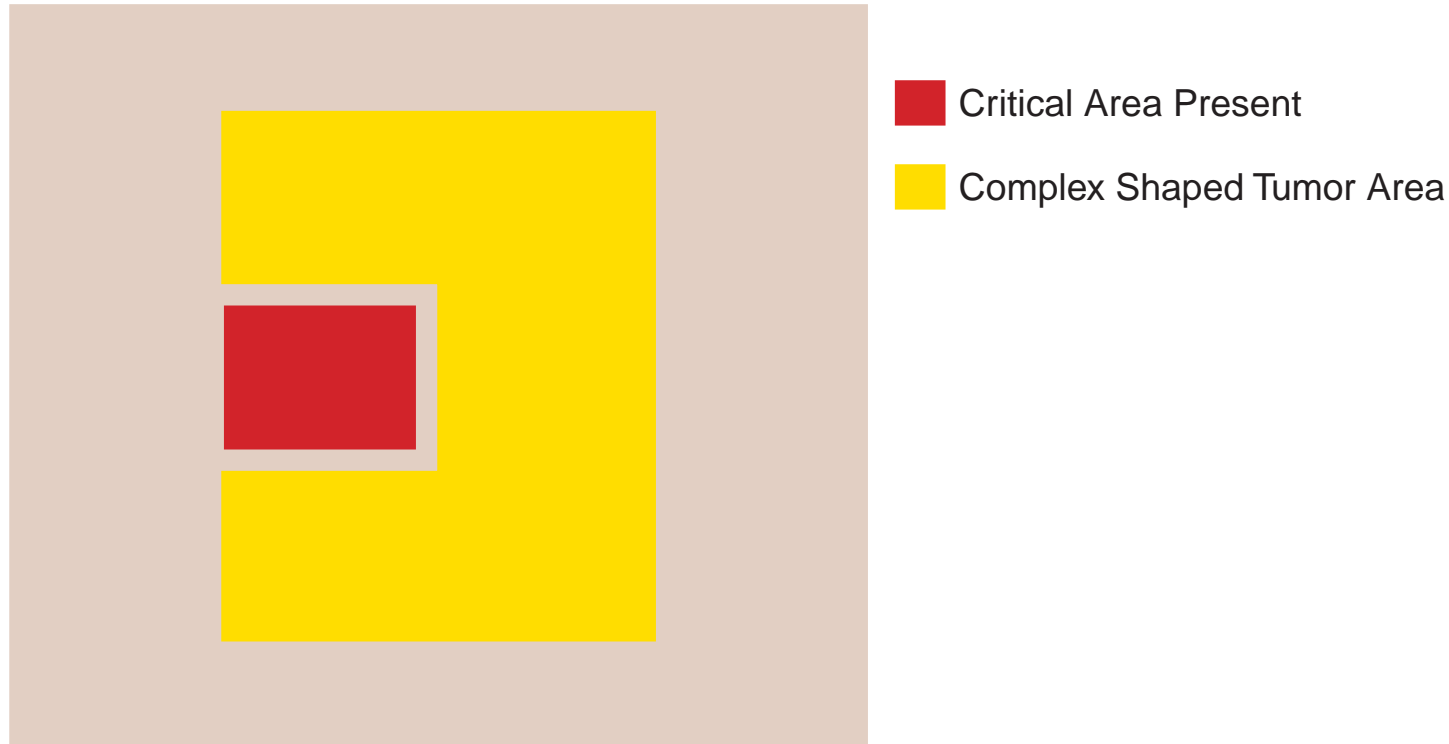
In conventional radiotherapy

- 3 to 7 beams of radiation
- radiation oncologist and physicist work together to determine a set of beam angles and beam intensities
- determined by manual “trial-and-error” process

Radiation Therapy

Overview

...Conventional Radiotherapy



With only a small number of beams, it is difficult/impossible to deliver required dose to tumor without impacting the critical area.

Radiation Therapy

Overview

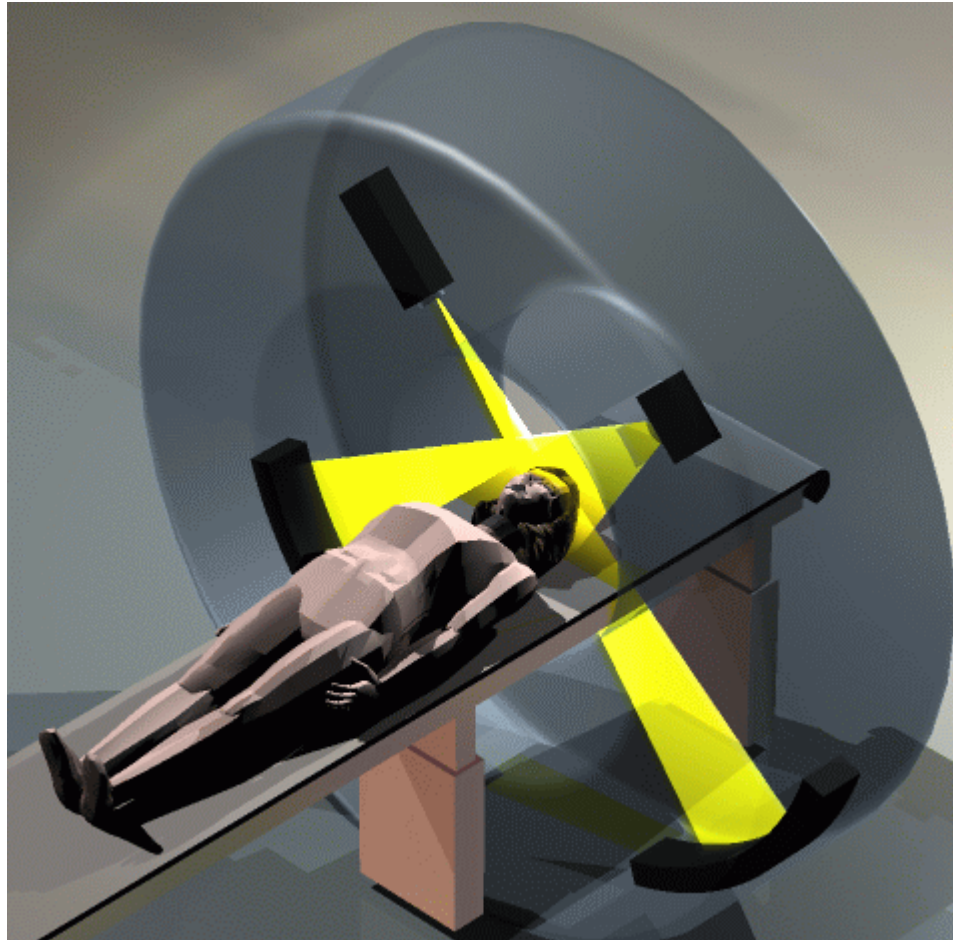
Recent Advances...

- More accurate map of tumor area
 - CT — Computed Tomography
 - MRI — Magnetic Resonance Imaging
- More accurate delivery of radiation
 - IMRT: Intensity Modulated Radiation Therapy
 - Tomotherapy

Radiation Therapy

Overview

...Recent Advances



Radiation Therapy

Overview

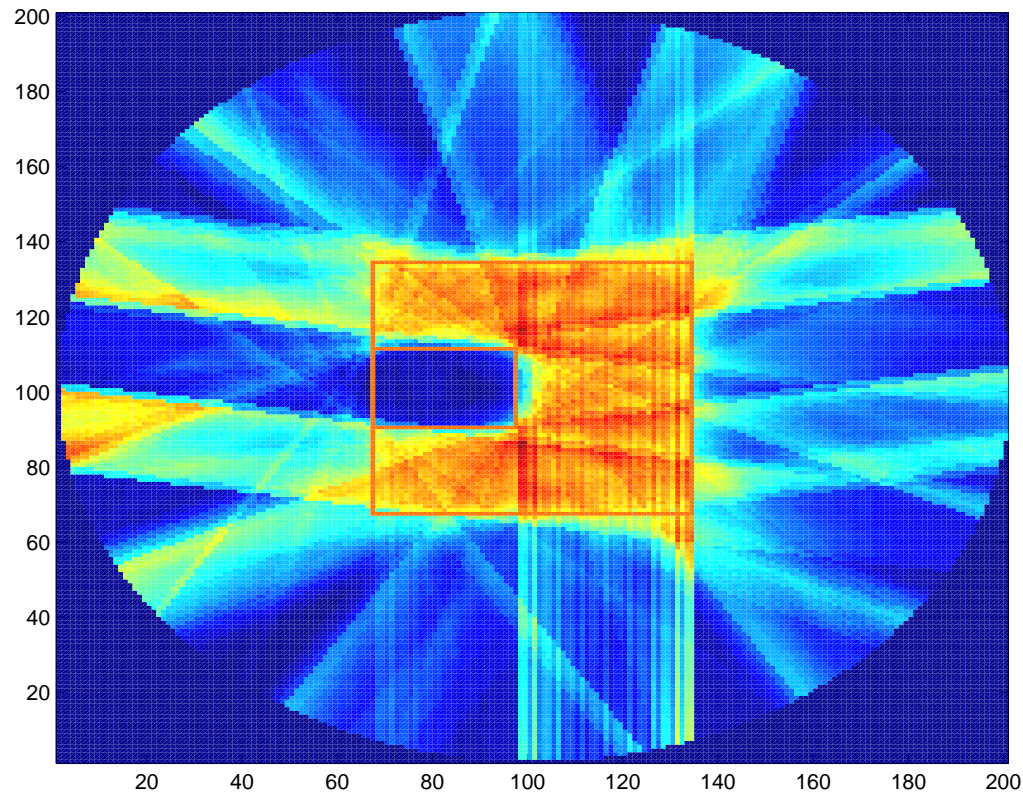
Formal Problem Statement...

- For a given tumor and given critical areas
- For a given set of possible beamlet origins and angles
- Determine the weight on each beamlet such that:
 - dosage over the tumor area will be at least a target level γ_L
 - dosage over the critical area will be at most a target level γ_U

Radiation Therapy

Overview

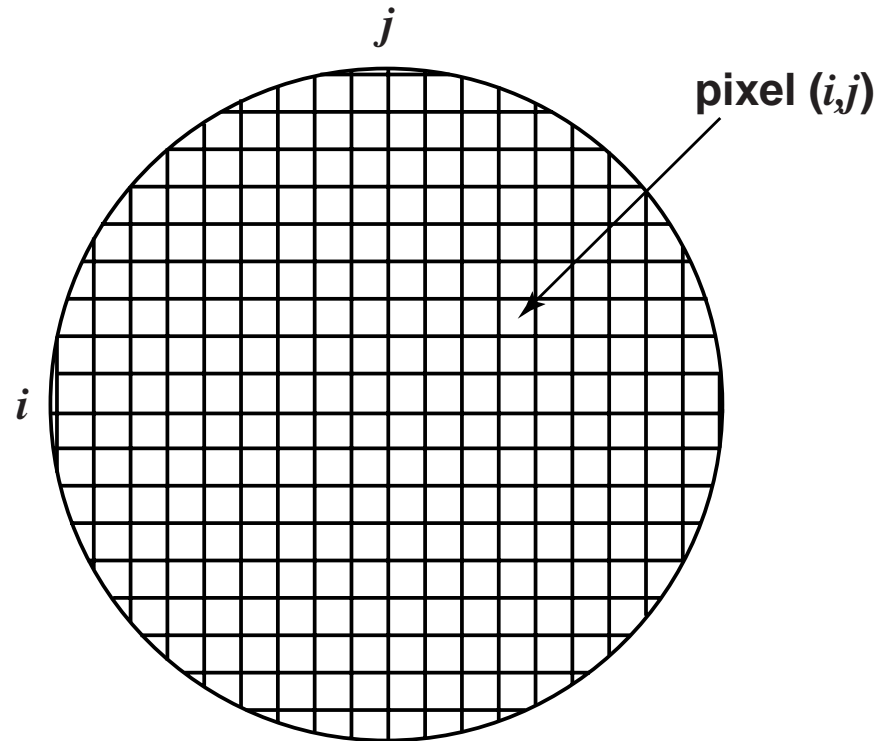
...Formal Problem Statement



Linear Optimization Models

Discretize the Space

Divide up region into a 2-dimensional (or 3-dimensional) grid of pixels



Linear Optimization Models

Create Beamlet Data

Create the beamlet data for each of $p = 1, \dots, n$ possible beamlets.

D^p is the matrix of unit doses delivered by beam p .

0	0	0	0	0.8	0.8	0	0
0	0	0	0.9	0.9	0	0	0
0	0	0	0.9	0.9	0	0	0
0	0	1.0	1.0	0	0	0	0
0	1.0	1.0	0	0	0	0	0

$D_{ij}^p =$ unit dose delivered to pixel (i, j) by beamlet p .

Linear Optimization Models

Dosage Equations

Decision variables $w = (w_1, \dots, w_n)$

w_p = intensity weight assigned to beamlet p ,

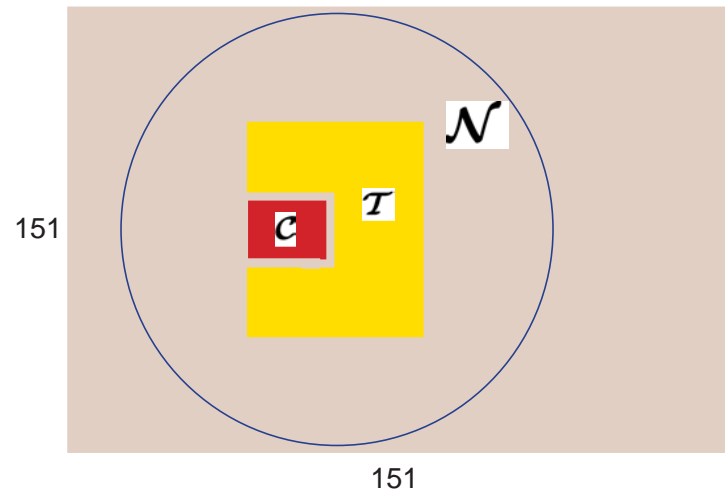
$p = 1, \dots, n.$

$$D_{i j} := \sum_{p=1}^n D_{i j}^p w_p$$

(“:=” denotes “by definition”)

$$D := \sum_{p=1}^n D^p w_p$$

is the matrix of the integral dose (total delivered dose)



\mathcal{T} is the target area

\mathcal{C} is the critical area

\mathcal{N} is normal tissue

$$\mathcal{S} := \mathcal{T} \cup \mathcal{C} \cup \mathcal{N}$$

Linear Optimization Models

Ideal Linear Model

$$\begin{aligned} & \text{minimize} && \sum_{(i,j) \in \mathcal{S}} D_{ij} \\ & \quad w, D && n \\ & \text{s.t.} && D_{ij} = \sum_{p=1}^n D_{ij}^p w_p && (i, j) \in \mathcal{S} \\ & && w \geq 0 \\ & && D_{ij} \geq \gamma_L && (i, j) \in \mathcal{T} \\ & && D_{ij} \leq \gamma_U && (i, j) \in \mathcal{C} \end{aligned}$$

Linear Optimization Models

Ideal Linear Model

$$\begin{aligned} & \underset{w, D}{\text{minimize}} && \sum_{(i,j) \in \mathcal{S}} D_{ij} \\ & \text{s.t.} && D_{ij} = \sum_{p=1}^n D_{ij}^p w_p && (i, j) \in \mathcal{S} \\ & && w \geq \mathbf{0} \\ & && D_{ij} \geq \gamma_L && (i, j) \in \mathcal{T} \\ & && D_{ij} \leq \gamma_U && (i, j) \in \mathcal{C} \end{aligned}$$

- Unfortunately, this model is typically infeasible.
- Cannot deliver dose to tumor without some harm to critical area(s).

Linear Optimization Models

Engineered Approaches

$$\begin{aligned} \text{minimize}_{w, D} \quad & \theta_{\mathcal{T}} \sum_{(i,j) \in \mathcal{T}} D_{ij} + \theta_{\mathcal{C}} \sum_{(i,j) \in \mathcal{C}} D_{ij} + \theta_{\mathcal{N}} \sum_{(i,j) \in \mathcal{N}} D_{ij} \\ \text{s.t.} \quad & D_{ij} = \sum_{p=1}^n D_{ij}^p w_p \quad (i, j) \in \mathcal{S} \\ & w \geq \mathbf{0} \\ & \gamma_{ij}^L \leq D_{ij} \leq \gamma_{ij}^U \quad (i, j) \in \mathcal{T} \\ & w_m \leq \frac{\alpha}{n} \sum_{p=1}^n w_p \quad m = 1, \dots, n \\ & \text{(typically use } \alpha = 5 \text{)} \end{aligned}$$

Linear Optimization Models

Engineered Approaches

Some other possible objective functions:

Let $(\text{Target})_{i,j}$ be the target prescribed dose to be delivered to pixel (i, j)

$$\begin{aligned} &\text{minimize}_{w,D} && \max_{(i,j) \in \mathcal{S}} |D_{i,j} - (\text{Target})_{i,j}| \\ &\text{s.t.} && D_{i,j} = \sum_{p=1}^n D_{i,j}^p w_p \quad (i, j) \in \mathcal{S} \\ &&& w \geq \mathbf{0} \end{aligned}$$

Linear Optimization Models

Engineered Approaches

This is the same as:

$$\text{minimize } \mu$$

w, D, μ

$$\text{s.t.} \quad -\mu \leq D_{i j} - (\text{Target})_{i j} \leq \mu \quad (i, j) \in \mathcal{S}$$

$$D_{i j} = \sum_{p=1}^n D_{i j}^p w_p \quad (i, j) \in \mathcal{S}$$

$$w \geq 0$$

Linear Optimization Models

Engineered Approaches

Here is another model:

$$\begin{aligned} &\underset{w, D}{\text{minimize}} && \sum_{(i,j) \in \mathcal{S}} |D_{i j} - (\text{Target})_{i j}| \\ &\text{s.t.} && D_{i j} = \sum_{p=1}^n D_{i j}^p w_p \quad (i, j) \in \mathcal{S} \\ &&& w \geq \mathbf{0} \end{aligned}$$

Linear Optimization Models

Engineered Approaches

This is the same as:

$$\text{minimize}_{w, D, \Delta} \sum_{(i,j) \in \mathcal{S}} \Delta_{i j}$$

$$\text{s.t.} \quad D_{i j} = \sum_{p=1}^n D_{i j}^p w_p \quad (i, j) \in \mathcal{S}$$

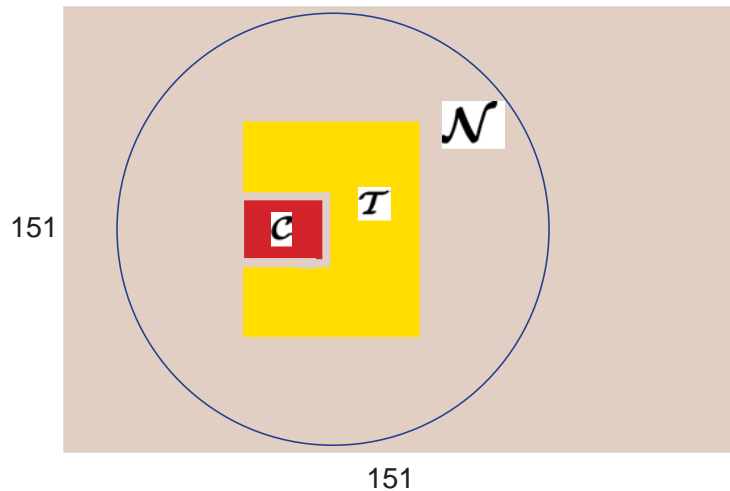
$$w \geq 0$$

$$-\Delta_{i j} \leq D_{i j} - (\text{Target})_{i j} \leq \Delta_{i j} \quad (i, j) \in \mathcal{S}$$

Computation

Base Case Model

Consider the “base case” example problem:



$$(\text{Target})_{i j} = 16, \quad (i, j) \in \mathcal{T}$$

$$(\text{Target})_{i j} = 0, \quad (i, j) \in \mathcal{C}$$

$$(\text{Target})_{i j} = 0, \quad (i, j) \in \mathcal{N}$$

Computation

Base Case Model

$$\text{minimize}_{w, D, \Delta} \quad 1 \cdot \sum_{(i,j) \in \mathcal{N}} \Delta_{i,j} + 100 \sum_{(i,j) \in \mathcal{C}} \Delta_{i,j} + 30 \sum_{(i,j) \in \mathcal{T}} \Delta_{i,j}$$

$$\text{s.t.} \quad D_{i,j} = \sum_{p=1}^n D_{i,j}^p w_p \quad (i,j) \in \mathcal{S}$$
$$w \geq 0$$

$$-\Delta_{i,j} \leq D_{i,j} - (\text{Target})_{i,j} \leq \Delta_{i,j} \quad (i,j) \in \mathcal{S}$$

Computation

Size of the Model

Dimensional Analysis...

$$\begin{aligned} \text{minimize}_{w, D, \Delta} \quad & 1 \cdot \sum_{(i,j) \in \mathcal{N}} \Delta_{ij} + 100 \sum_{(i,j) \in \mathcal{C}} \Delta_{ij} + 30 \sum_{(i,j) \in \mathcal{T}} \Delta_{ij} \\ \text{s.t.} \quad & D_{ij} = \sum_{p=1}^n D_{ij}^p w_p \quad (i,j) \in \mathcal{S} \\ & w \geq 0 \\ & -\Delta_{ij} \leq D_{ij} - (\text{Target})_{ij} \leq \Delta_{ij} \quad (i,j) \in \mathcal{S} \end{aligned}$$

Dimensional Analysis:

number of pixels = 31,397 ($\approx \pi * 100^2$)

number of beamlets = 564 (n)

$|\mathcal{T}| = 3,859$; $|\mathcal{C}| = 630$; $|\mathcal{N}| = 26,908$

$|\mathcal{S}| = 31,397$

Computation

Size of the Model

...Dimensional Analysis...

$$\begin{aligned} \text{minimize}_{w,D,\Delta} \quad & 1 \cdot \sum_{(i,j) \in \mathcal{N}} \Delta_{ij} + 100 \sum_{(i,j) \in \mathcal{C}} \Delta_{ij} + 30 \sum_{(i,j) \in \mathcal{T}} \Delta_{ij} \\ \text{s.t.} \quad & D_{ij} = \sum_{p=1}^n D_{ij}^p w_p \quad (i,j) \in \mathcal{S} \\ & w \geq 0 \\ & -\Delta_{ij} \leq D_{ij} - (\text{Target})_{ij} \leq \Delta_{ij} \quad (i,j) \in \mathcal{S} \end{aligned}$$

Decision Variables	Number
D_{ij}	31,397
w	564
Δ_{ij}	31,397
Total	63,358

Computation

Size of the Model

...Dimensional Analysis

$$\begin{aligned} \text{minimize}_{w,D,\Delta} \quad & 1 \cdot \sum_{(i,j) \in \mathcal{N}} \Delta_{i,j} + 100 \sum_{(i,j) \in \mathcal{C}} \Delta_{i,j} + 30 \sum_{(i,j) \in \mathcal{T}} \Delta_{i,j} \\ \text{s.t.} \quad & D_{i,j} = \sum_{p=1}^n D_{i,j}^p w_p \quad (i,j) \in \mathcal{S} \\ & w \geq 0 \\ & -\Delta_{i,j} \leq D_{i,j} - (\text{Target})_{i,j} \leq \Delta_{i,j} \quad (i,j) \in \mathcal{S} \end{aligned}$$

Computation

Size of the Model

Number of Constraints

Simple Variables Upper/Lower Bounds	Number
$w \geq 0$	564
Total	564

Other Constraints*	Number
$D_{ij} =$	31,397
$\leq D_{ij} - (\text{Target})_{ij} \leq$	62,794
Total	94,191

*We usually exclude simple variable upper/lower bounds when counting constraints.

Computation

Size of the Model

Summary

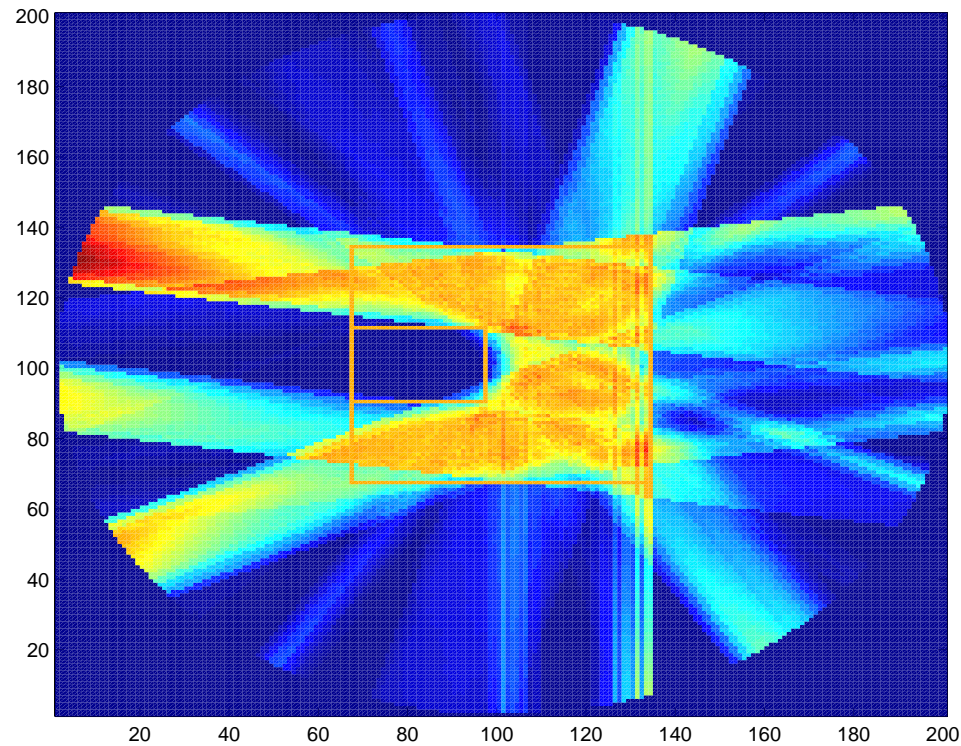
Variables	Constraints*
63, 358	94, 191

*Excludes variable upper/lower bounds.

Computation

Base Case Model

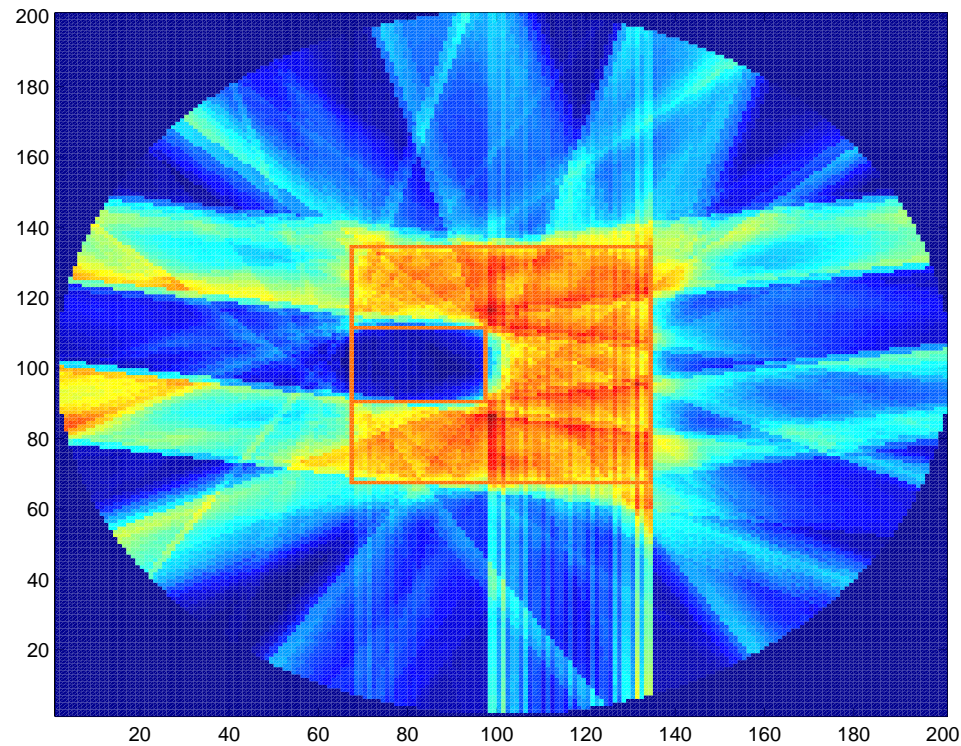
Optimal Solution



Base Case Model Solution

Computation

Another Model Solution

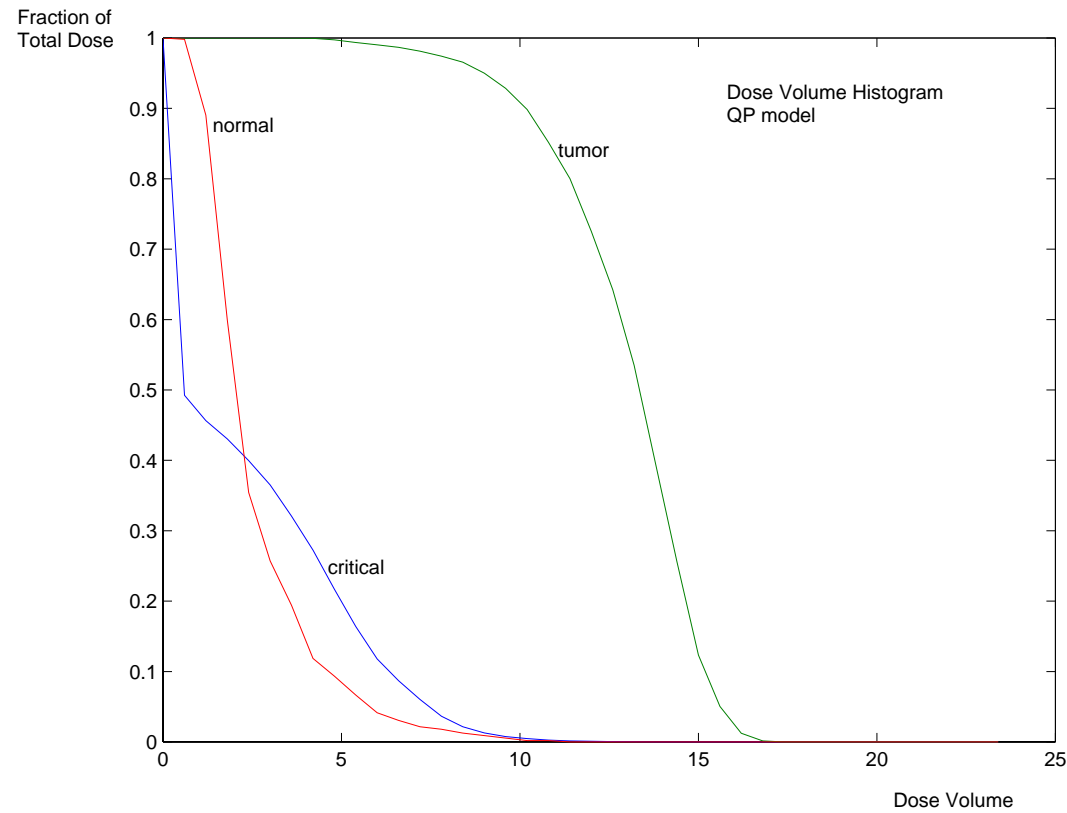


Solution of a nonlinear model.

Computation

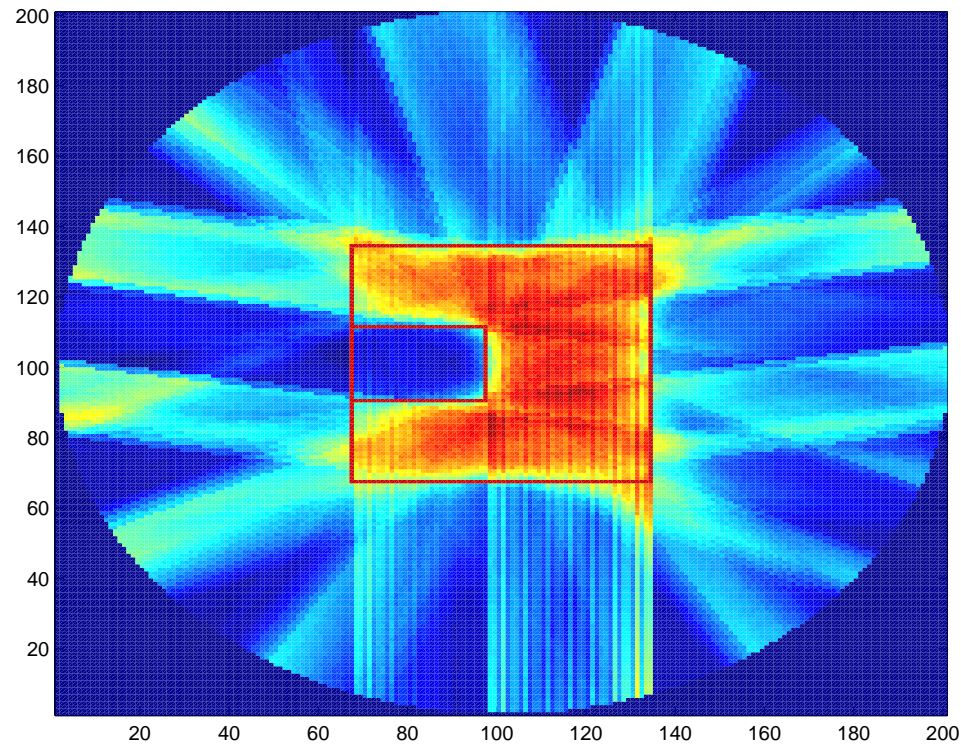
Dose Histogram

of Solution



Computation

Another Model Solution



Solution of a nonlinear model, where $\theta_{\mathcal{N}} = \theta_{\mathcal{C}} = \theta_{\mathcal{T}} = 1$.

Computation

Computational Issues

Software/Algorithms

- Software codes:
 - CPLEX simplex (pivoting method)
 - CPLEX barrier
 - LOQO
- Algorithms:
 - Simplex method (“pivoting” method)
 - Interior-point method (IPM) (“barrier” method)

Computation

Computational Issues

Counting Iterations

- Iteration Counts:
 - Number of pivots for simplex method
 - Number of Newton steps for IPM

Computation

Computational Issues

Issues in Running Times

- Running time will be affected by:
 - number of constraints
 - number of variables
 - software code
 - type of algorithm (simplex or IPM)
 - properties of linear algebra systems involved
 - * density/patterns of nonzeros of matrix systems to be solved
 - other problem characteristics specific to problem
 - idiosyncratic influences
 - pre-processing heuristics

Computation

Base Case

No Pre-Processing

- Base Case Model
- No Pre-Processing

Code	Algorithm	Iterations	Running Time	
			CPU (sec)	Wall (minutes)
CPLEX	Simplex	183,530	440	250
CPLEX	Barrier	49	13	37

1. The simplex algorithm is designed to handle variables with lower bounds and upper bounds:

$$\begin{aligned} \min_{\mathbf{x}} \quad & \mathbf{c}^T \mathbf{x} \\ & \mathbf{A}\mathbf{x} = \mathbf{b} \\ & \underline{\ell} \leq \mathbf{x} \leq \underline{u} \end{aligned}$$

where $\ell_j = -\infty$ and/or $u_j = +\infty$ is allowed.

2. We say x_j has no bounds if $\ell_j = -\infty$ and $u_j = +\infty$. Otherwise x_j is a bounded variable.

Computation

Some Generic Rules

$$\begin{aligned} \min_x \quad & c^T x \\ & Ax = b \\ & \ell \leq x \leq u \end{aligned}$$

3. For the simplex method, the work per pivot generally depends on the number of nonzeros in A .
4. If A is very sparse (its density of nonzero elements is low), then the work per pivot will be low.
5. The number of simplex pivots in a “good” model is roughly between m and $10n$.

$$\begin{aligned} \min_x \quad & c^T x \\ & Ax = b \\ & \ell \leq x \leq u \end{aligned}$$

5. The work per iteration of an interior-point method generally depends on the structure of the matrix

$$K = \begin{pmatrix} I & A^T \\ A & \mathbf{0} \end{pmatrix}.$$

$$K = \begin{pmatrix} I & A^T \\ A & \mathbf{0} \end{pmatrix}.$$

6. The structure of K is often (but not always) related to the structure of the matrix AA^T because the following two matrices are “similar”:

$$K = \begin{pmatrix} I & A^T \\ A & \mathbf{0} \end{pmatrix} \quad P = \begin{pmatrix} I & A^T \\ \mathbf{0} & -AA^T \end{pmatrix}.$$

7. The number of interior-point method iterations is typically **25–80** (*independent of m and/or n*).

Computation

Pre-Processing

Heuristics...

Pre-Processing Heuristics in Commercial-Grade Software

- Designed to Eliminate Constraints and/or Variables
- Example:

$$-5x \quad +3y \quad +z \quad = 17$$

$$0 \leq x \leq 4 \quad 0 \leq y \leq 2 \quad 10 \leq z \leq 40$$

Pre-Processing

Computation

...Heuristics...

- Example:

$$-5x \quad +3y \quad +z \quad = 17$$

$$0 \leq x \leq 4 \quad 0 \leq y \leq 2 \quad 10 \leq z \leq 40$$

- $z = 17 + 5x - 3y \geq 17 + 5(0) - 3(2) = 11 \geq 10$
- $z = 17 + 5x - 3y \leq 17 + 5(4) - 3(0) = 37 \leq 40$
- Therefore we can eliminate the bounds on z
- Therefore we can treat z as a free variable
- Therefore we can eliminate z from our model altogether.

Computation

Pre-Processing

...Heuristics

- Base Case Model
- With Pre-Processing

Code	Algorithm	Iterations	Running Time	
			CPU (sec)	Wall (minutes)
CPLEX	Simplex	18,428	4.3	4
CPLEX	Barrier	16	130	133

Computation

Equivalent Formulation

“Small” Model...

Equivalent Formulation: (eliminate D_{ij})

“Small” Model:

$$\begin{aligned} \text{minimize}_{w, \Delta} \quad & 1 \cdot \sum_{(i,j) \in \mathcal{N}} \Delta_{ij} + 100 \sum_{(i,j) \in \mathcal{C}} \Delta_{ij} + 30 \sum_{(i,j) \in \mathcal{T}} \Delta_{ij} \\ \text{s.t.} \quad & -\Delta_{ij} \leq \sum_{p=1}^n D_{ij}^p w_p - (\text{Target})_{ij} \leq \Delta_{ij} \quad (i,j) \in \mathcal{S} \\ & w \geq 0 \end{aligned}$$

Computation

Equivalent Formulation

...“Small” Model...

	Base Case Model	Small Model
Variables	63, 358	31, 961
Constraints*	94, 191	62, 794

*always excludes simple variable upper/lower bounds

Computation

Equivalent Formulation

...“Small” Model

- Small Model

Code	Algorithm	Iterations	Running Time	
			CPU (sec)	Wall (minutes)
CPLEX	Simplex	171,656	390	216
CPLEX	Barrier	57	80	31

Computation

Comparisons

Code	Algorithm	Model	Running Time
			Wall
			(minutes)
CPLEX	Simplex	Base Case	250
		Pre-Processed	4
		Small Model	216
CPLEX	Barrier	Base Case	37
		Pre-Processed	133
		Small Model	31

Nonlinear Optimization

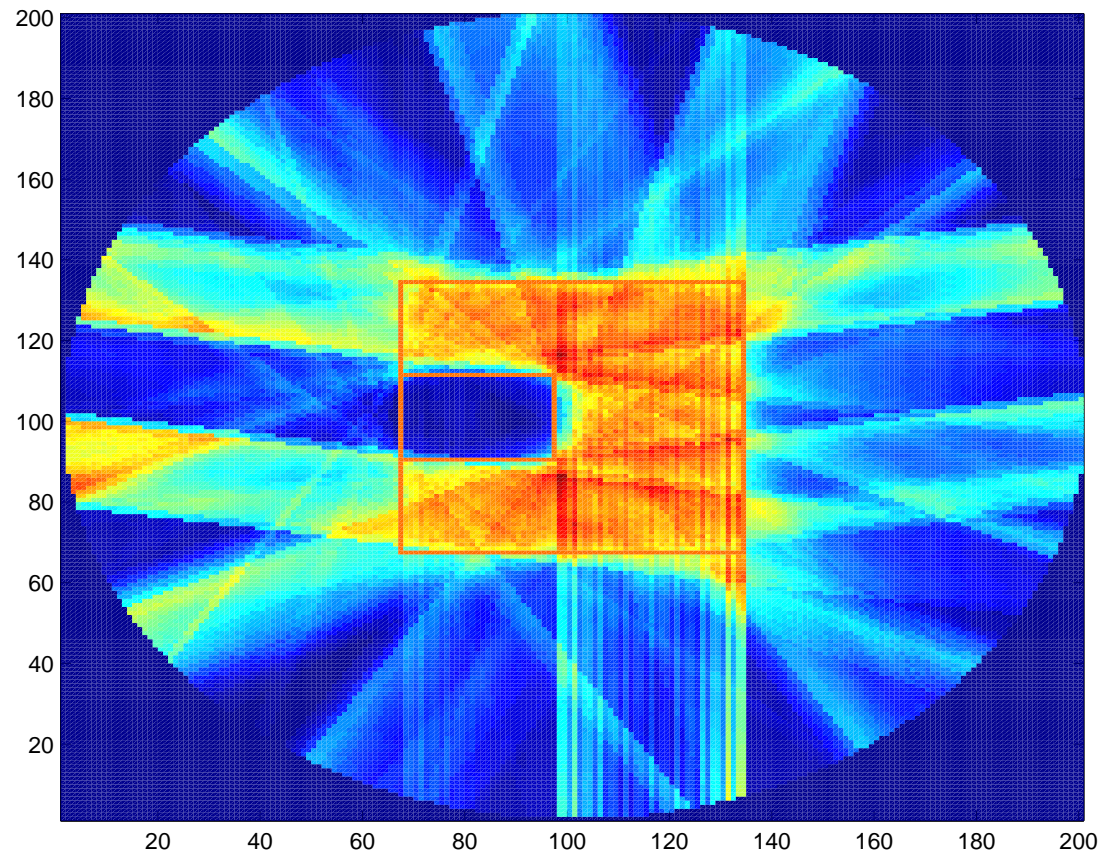
Quadratic Model

$$\begin{aligned} QP: \text{ minimize}_{w,D} \quad & 1 \cdot \sum_{(i,j) \in \mathcal{N}} [D_{ij} - \text{Target}_{ij}]^2 \\ & + 100 \sum_{(i,j) \in \mathcal{C}} [D_{ij} - \text{Target}_{ij}]^2 \\ & + 30 \sum_{(i,j) \in \mathcal{T}} [D_{ij} - \text{Target}_{ij}]^2 \\ \text{ s.t.} \quad & D_{ij} = \sum_{p=1}^n D_{ij}^p w_p \quad (i,j) \in \mathcal{S} \\ & w \geq 0 \end{aligned}$$

Nonlinear Optimization

Quadratic Model

Quadratic Model Output



Nonlinear Optimization

Quadratic Model

Computational Results

Model	Code	Algorithm	Iterations	Running Time
				CPU (sec)
Base Case QP Model	LOQO	Barrier	31	82.7
Small QP Model	LOQO	Barrier	32	149.0

Mixed Integer Optimization

Limiting the Number of Beamlets

$$\begin{aligned} \text{minimize}_{w, D, \Delta} \quad & 1 \cdot \sum_{(i,j) \in \mathcal{N}} \Delta_{ij} + 100 \sum_{(i,j) \in \mathcal{C}} \Delta_{ij} + 30 \sum_{(i,j) \in \mathcal{T}} \Delta_{ij} \\ \text{s.t.} \quad & D_{ij} = \sum_{p=1}^n D_{ij}^p w_p & (i,j) \in \mathcal{S} \\ & w \geq 0 \\ & -\Delta_{ij} \leq D_{ij} - (\text{Target})_{ij} \leq \Delta_{ij} & (i,j) \in \mathcal{S} \\ & w_p \leq 100 y_p & p = 1, \dots, n \\ & y_p \in \{0, 1\} & p = 1, \dots, n \\ & \sum_{p=1}^n y_p \leq 15. \end{aligned}$$

Mixed Integer Optimization

Computation

CPLEX MIP Solver

MIP Gap (%)	Simplex Pivots	Running Time	
		CPU (seconds)	Wall (minutes)
20	11,646	7	4
15	11,646	7	4
12	11,646	5	4
10	14,538	9	6
7	14,538	7	6
5	14,538	10	6
4	14,538	7	6
3	14,538	5	6
2	3,655,445	1,700	25.3 hours

Modifications of the Model

Partial Volume Constraints

Partial Volume Constraints:

“No more than 20% of the critical region can exceed a dose of $30G_y$.”

“No more than 5% of the critical region can exceed a dose of $50G_y$.”

Modifications of the Model

Partial Volume Constraints

Approach #1 (Integer Programming Model)

Let M be a very large number,

$$D_{i j} \leq 30 + M \cdot y_{i j}, \quad y_{i j} \in \{0, 1\}, \quad (i j) \in \mathcal{C}$$

$$D_{i j} \leq 50 + M \cdot z_{i j}, \quad z_{i j} \in \{0, 1\}, \quad (i j) \in \mathcal{C}$$

$$\sum_{(i j) \in \mathcal{C}} y_{i j} \leq |\mathcal{C}| \times 0.20$$

$$\sum_{(i j) \in \mathcal{C}} z_{i j} \leq |\mathcal{C}| \times 0.05$$

Modifications of the Model

Partial Volume Constraints

Approach #2 (Error Function Approach)

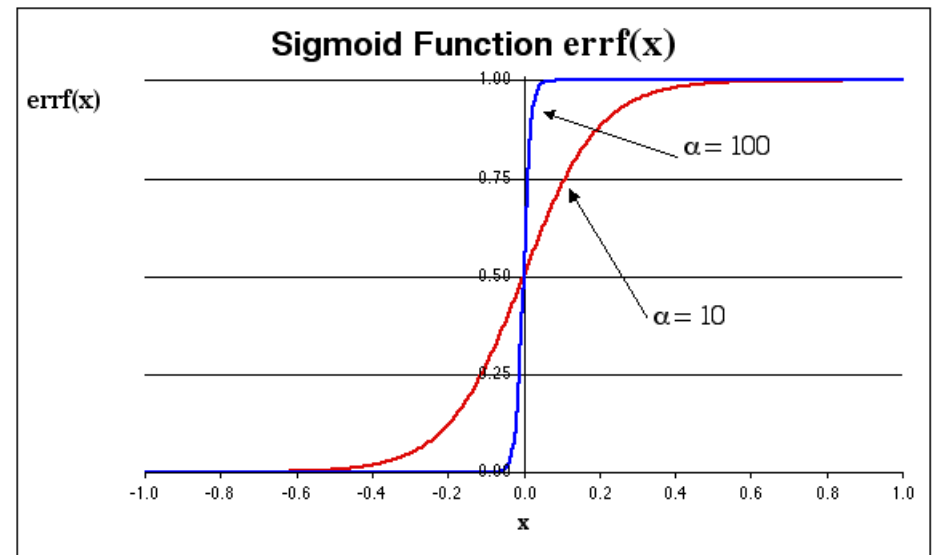
The *error function*, or *sigmoid function*, is of the form:

$$\text{err } f(x) = \frac{1}{1 + e^{-\alpha x}}$$

$$\text{err } f(x) = \frac{1}{2} \text{ at } x = 0$$

$$\text{err } f(x) \rightarrow 1 \text{ as } x \rightarrow \infty$$

$$\text{err } f(x) \rightarrow 0 \text{ as } x \rightarrow -\infty$$



Modifications of the Model

Partial Volume Constraints

Instead of integer variables, we use

$$\sum_{(i,j) \in \mathcal{C}} \text{err } f(D_{ij} - 30) \leq |\mathcal{C}| \times 0.20$$

$$\sum_{(i,j) \in \mathcal{C}} \text{err } f(D_{ij} - 50) \leq |\mathcal{C}| \times 0.05$$

Looking Ahead

Modeling Languages

Used in the Course

- Modeling languages and software used in the course
 - OPL Studio
 - * linear and mixed-integer programming
 - * solver is CPLEX simplex and/or CPLEX barrier
 - * first half of course
 - AMPL
 - * linear and nonlinear programming
 - * solver is LOQO
 - * second half of course

Modeling Tools

Looking Ahead

and Issues

- “Column Generation” (week 3)
 - generates new decision variables “on the fly”
- Exact optimization and exact feasibility
 - in models
 - in algorithms
- Computational Issues in LP (next lecture)
 - simplex method with upper/lower bounds
 - methods for updating the basis inverse