CSE543T: Algorithms for Nonlinear Optimization

Yixin Chen, PhD
Professor
Department of Computer Science & Engineering
Washington University in St Louis

Nonlinear Programming (NLP) problem

Minimize f(x)

(optional) Subject to h(x) =0

g(x) <= 0

Characteristics:

– variable space *X:* continuous, discrete, mixed, *curves*

- Decision variables, control variables, system inputs

– function properties

- Closed form, evaluation process, stochastic, …

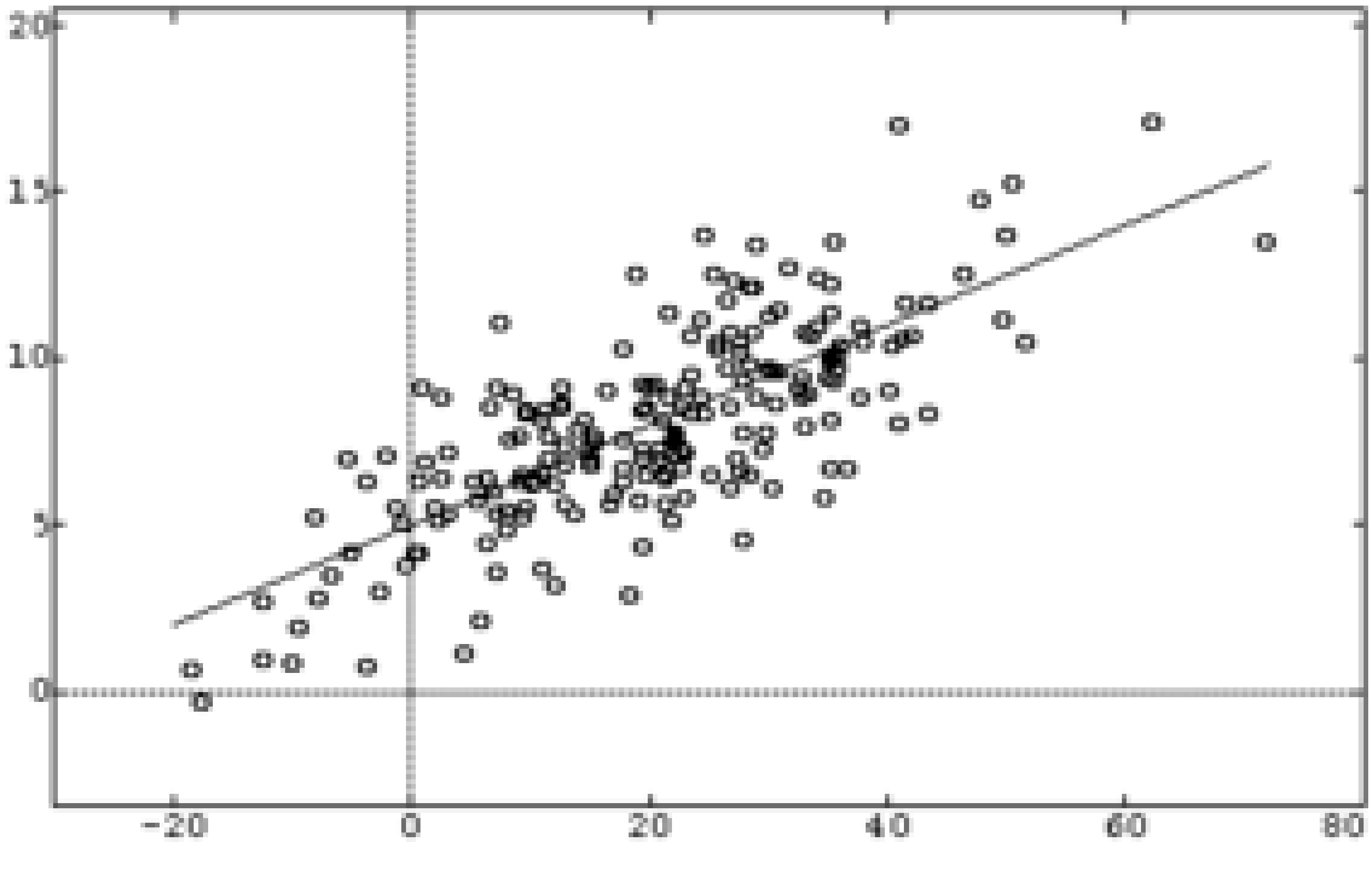
Why non-linear?

- Deal or no deal? You can choose
  - Take \$2M and go home; Or,
  - Deal: could get \$0.1 or \$5M
  - What will you choose?

Why study nonlinear optimization?

- Optimization is everywhere
  - Constraints are everywhere
  - The world is nonlinear
  - Need to make balanced decisions
- Plenty of applications
  - Investment, networking design, machine learning (NN, HMM, CRF), data mining, sensors, structural design, bioinformatics, medical treatment planning, games
- Nonlinear optimization is difficult
  - Curse of dimensionality, among many other reasons

Example: Linear Regression

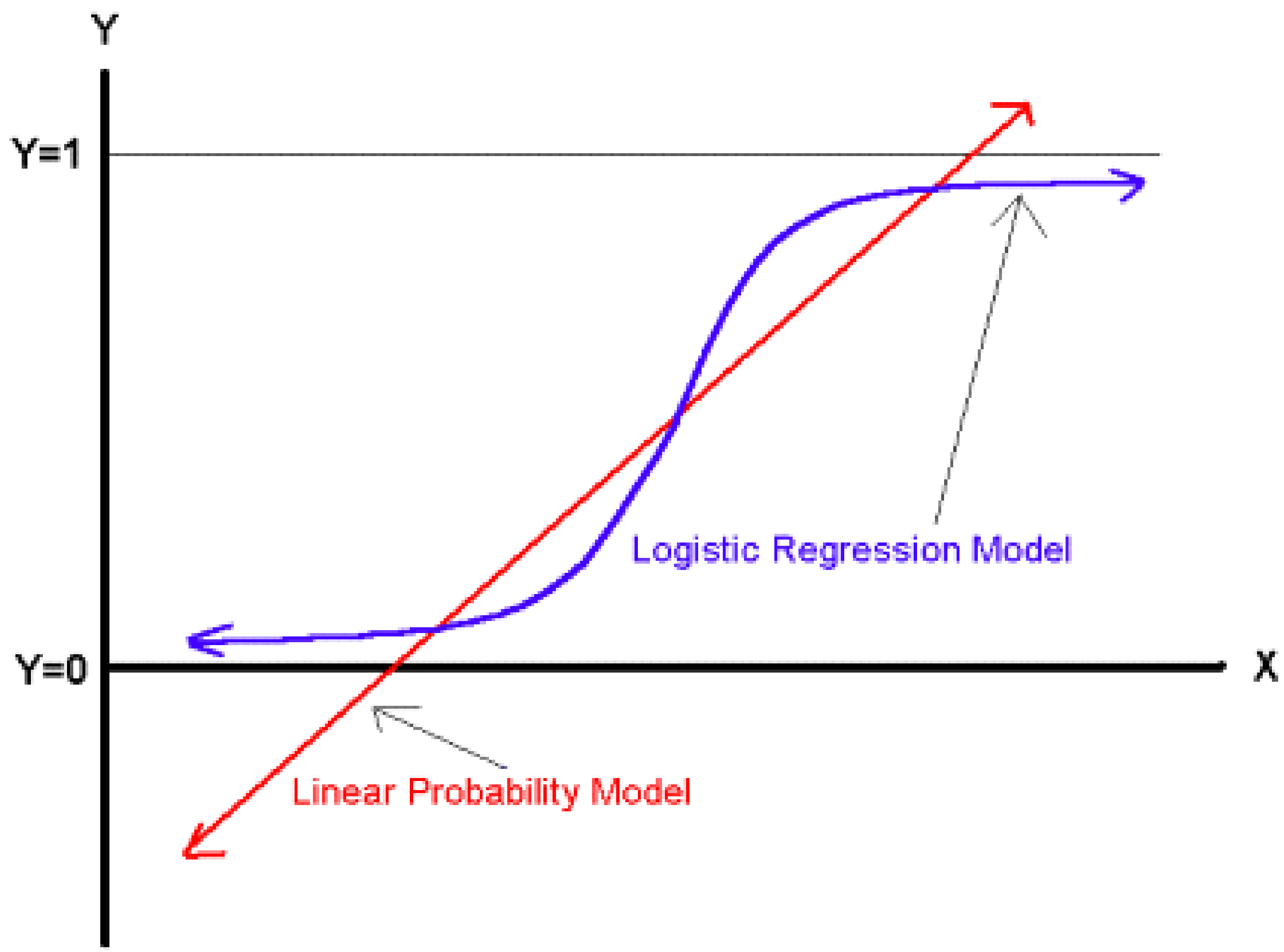


Logistic Regression

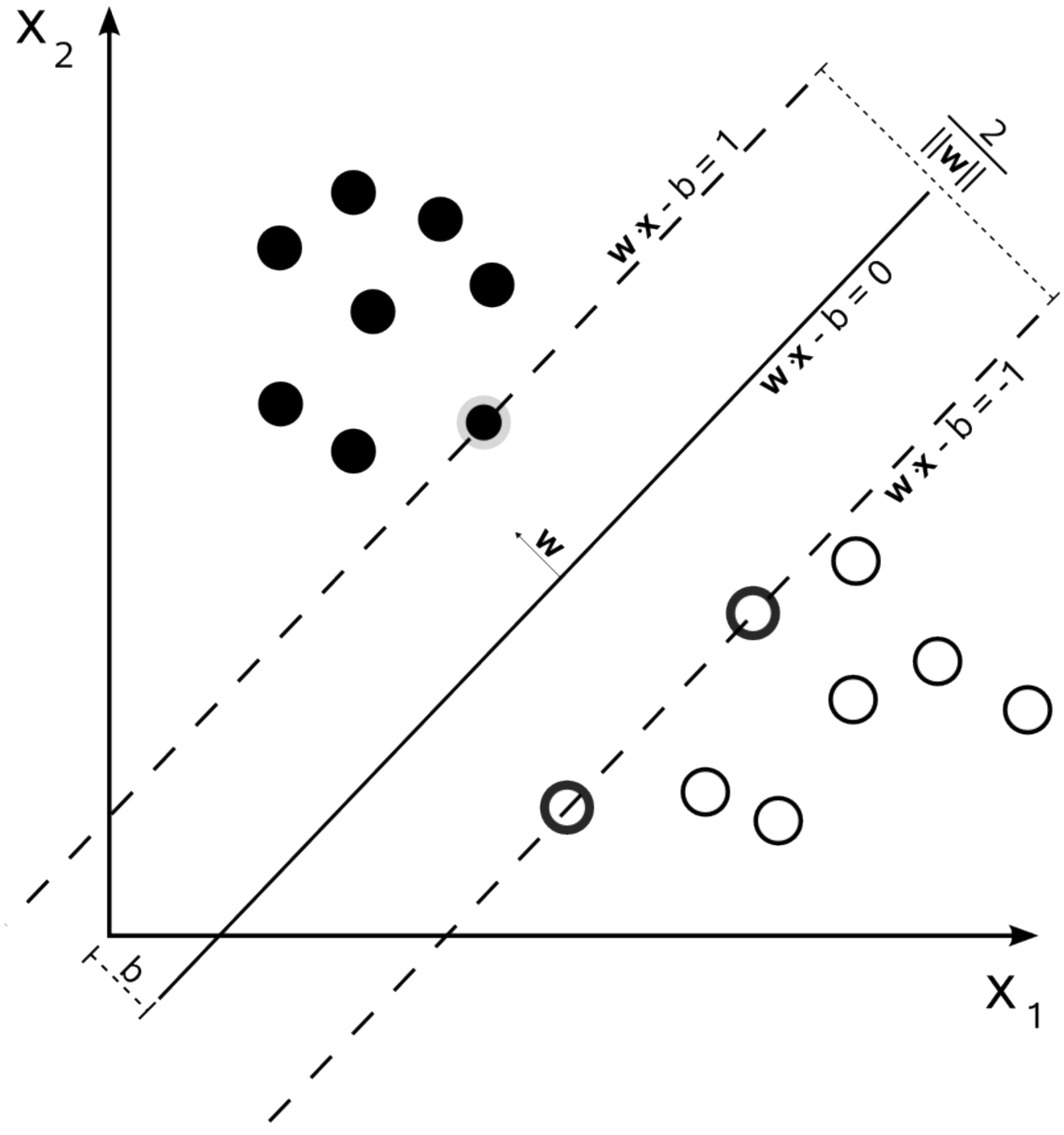
- Y=0 or 1 (e.g. stroke or not)
- Where nonlinear functions clearly make more sense
- $P(Y=1) = 1/(1+\exp(-\mathbf{w}^T\mathbf{x}))$
- Find **w** to maximize

$$\prod_{i=1..n} P(Y_i=1)^{Yi} (1-P(Y_i=1))^{(1-Yi)}$$

Comparing the LP and Logit Models

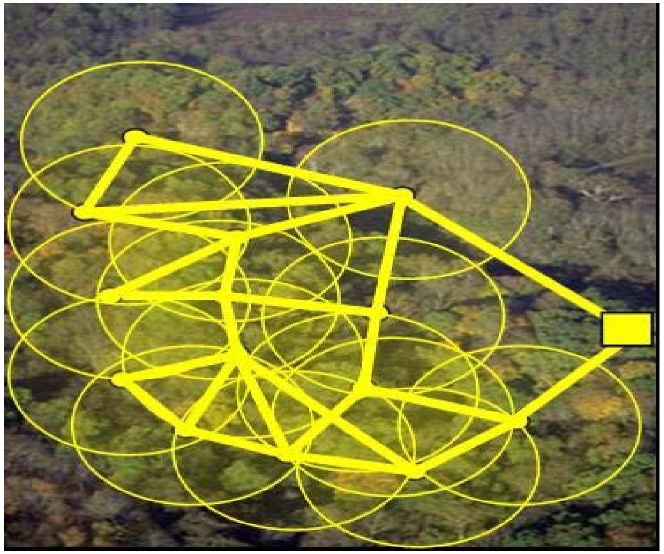


SVM



- Minimize
  $\|\mathbf{w}\|$
- subject to:
  $y_i(\mathbf{w}\ \mathbf{x}_i - b) \geq 1$

Sensor network optimization

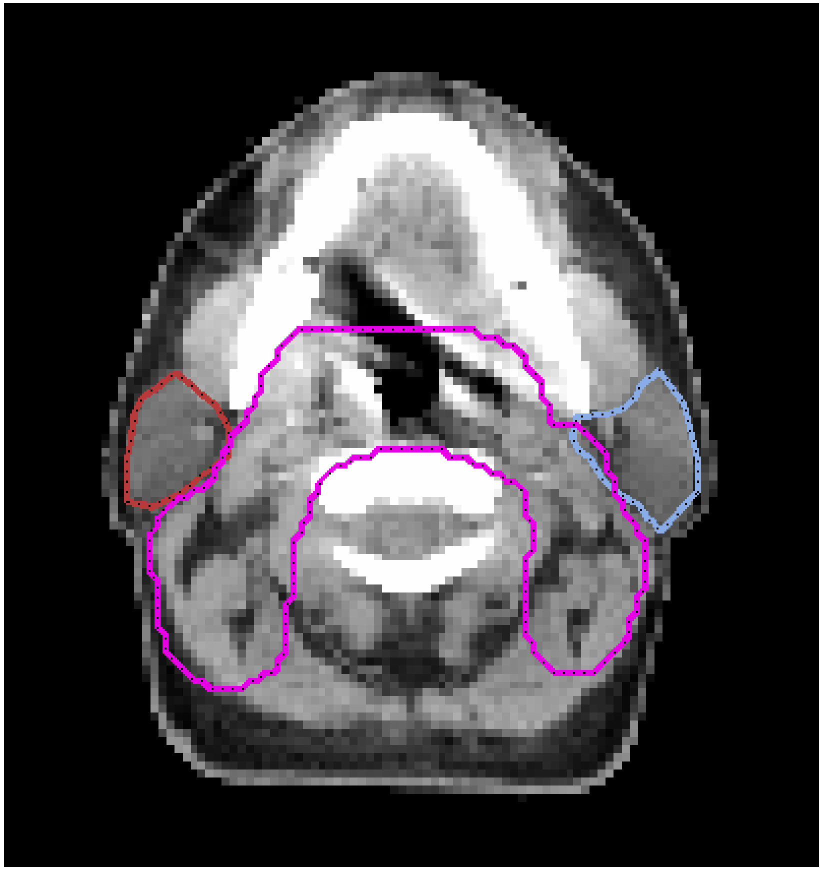


- Variables
  - Number of sensors
  - Locations of sensors
- Functions
  - Minimize the maximum false alarming rate
  - Subject to minimum detection probability > 95%
  - Highly nonlinear functions
    - Based on a Gaussian model and voting procedure
    - Computed by a Monte-Carlo simulation

Sodoku

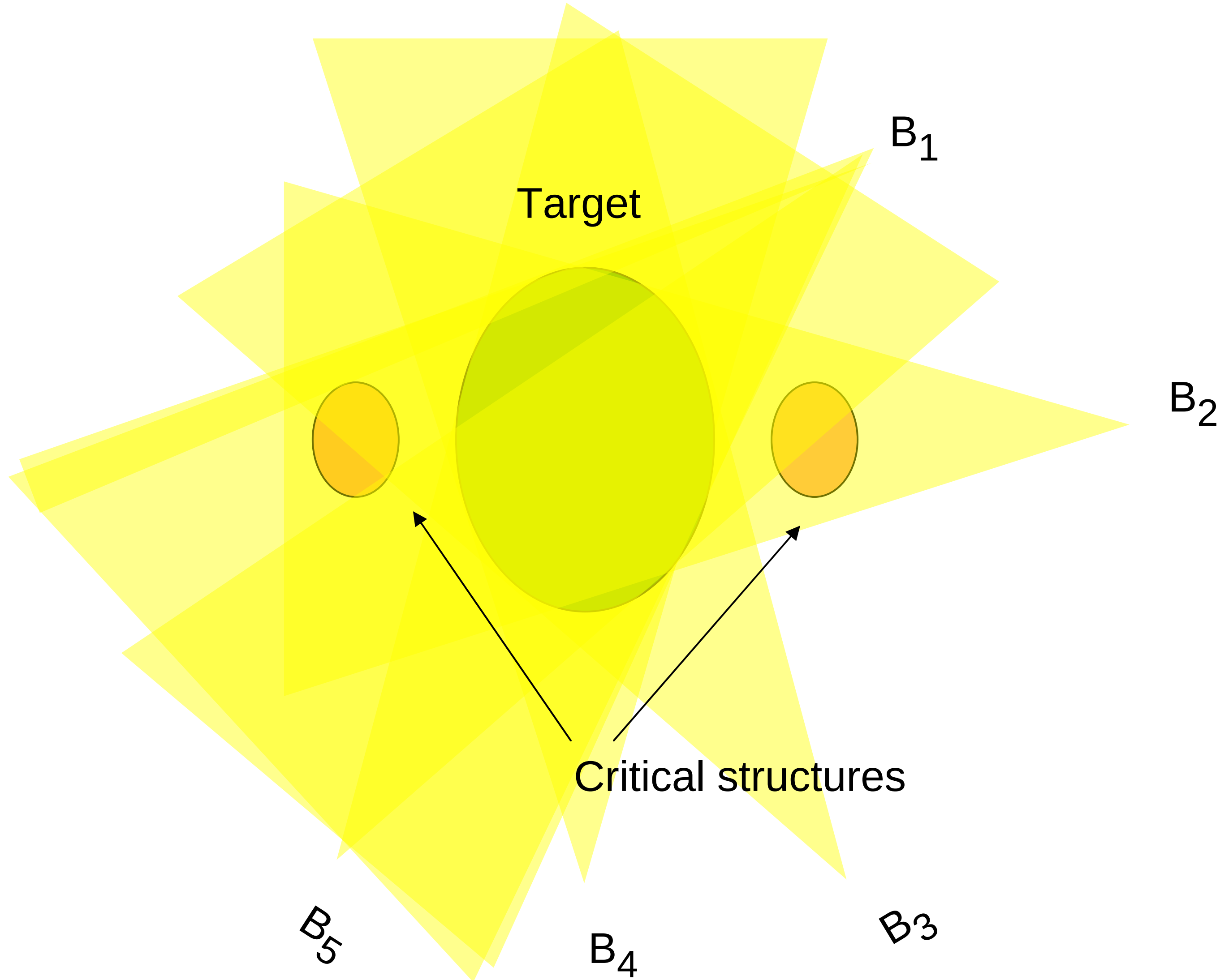
9			8		2			6
			9		3			
3		7				4		8
		2	5		7	8		
	4						3	
		6	1		4	2		
2		8				1		3
			3		9			
6			4		8			2

Application: Radiation Oncology Planning

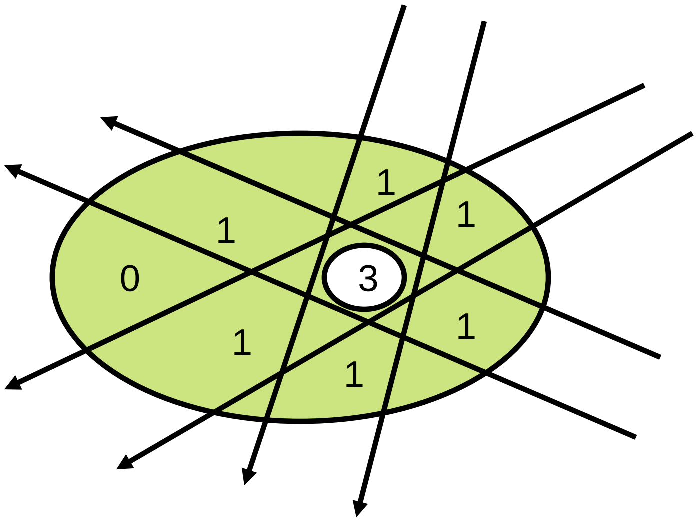


Example CT image of head and neck tumor

Intensity-modulated radiation therapy (IMRT)

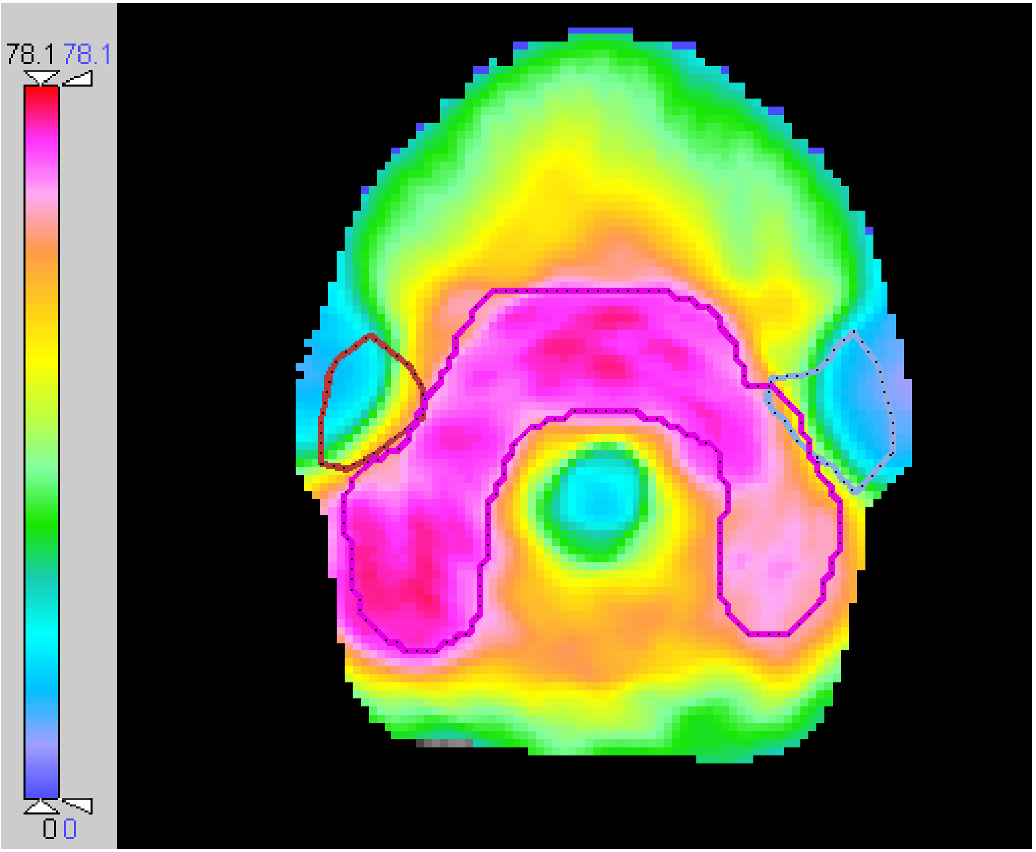


Field superposition: a basic principle of treatment planning



Schematic of patient cross section
and three incident unmodulated beams

Sample dose distribution resulting from IMRT beamlet optimization



Optimization quality: critical for survival chance of patients

Mathematical control variables

- Fixed:
  - Numbers of beams
  - Angles of incidence of beams
- Control variable
  - Fluence (amount of radiation) of each beamlets

Example Plan Acceptance Criteria

TABLE 2-3. BASIC PLAN ACCEPTANCE CRITERIA FOR CASES OF HEAD AND NECK CANCER

		Priority*
PTV-GTV	>95% volume received Rx dose >99% volume received 93% of Rx dose	2
PTV-CTV	>95% volume received Rx dose >99% of volume received 93% of Rx dose	2
Spinal cord	Maximum <45 Gy, or no more than 1 cc of volume >45 Gy	1
Brainstem	Maximum <54 Gy, or no more than 1% of volume >54 Gy	1
Optic structures	Maximum <54 Gy	2
Eye/retina	Maximum <50 Gy	3
Parotid glands	Mean <26 Gy, or 50% of volume <30 Gy	3
Mandible	Maximum <70 Gy, or no more than 1 cc of volume >70 Gy	3

PTV, planning target volume; GTV, gross tumor volume; CTV, clinical target volume; Rx, treatment.
* 1 is the highest priority.

From: Chao, C. (2005). *Practical Essentials of Intensity Modulated Radiation Therapy*, 2nd ed.

Formulation of Doses for Each Voxel

$$\mathbf{d} = \mathbf{A}\mathbf{w}$$

$$d_i = \sum_{j=1}^{\text{beamlets}} A_{i,j} w_j ,$$

The intensity for each voxel (unit of tissue) is the summation of contributing beamlets

Formulation: Constrained Nonlinear Optimization

$$\min \sum_{k=1}^{\#\text{ of structures}} c_k \frac{1}{N_k} \sum_{i=1}^{\text{N (\# of voxels)}} (d_i^k - d_{i,\text{prescribed}}^k)^2$$

$$\text{w.r.t.} \quad \mathbf{w}$$

with constraints

$$\max \ \mathbf{d}_{\text{spinal cord}} < 45 \text{ Gy}$$

$$\max \ \mathbf{d}_{\text{brain stem}} < 55 \text{ Gy}$$

$$\max \ \mathbf{d}_{\text{anywhere}} < 80 \text{ Gy}$$

• • • • • • • •

(Localized constraints, one for each region)

- Hundreds of thousands of variables
- Nonlinear constraints
- Very difficult to solve (has been studied for long)

Goals of the course

- Introduce the theory and algorithms for nonlinear optimization
  - One step further to the back of the blackbox
  - Able to hack solvers
  - But not too much gory details of mathematics
- Hands-on experiences with optimization packages
  - Know how to choose solvers
  - Understanding of the characteristics of problems/solvers
- Mathematical modeling
  - Modeling languages
  - CSE applications
- Help solve problems in your own research