Nonlinear optimization: Methods and applications

Anders Forsgren

Optimization and Systems Theory
Department of Mathematics
Royal Institute of Technology (KTH)
Stockholm, Sweden

Bridging the Gap
University of Birmingham
June 6, 2008

Outline

1. Background
   - Nonlinear optimization
2. Methods
   - Interior methods
3. Applications
   - Optimization of radiation therapy
   - Telecommunications optimization
4. On bridging the gap

Outline

1. Background
   - Nonlinear optimization
2. Methods
   - Interior methods
3. Applications
   - Optimization of radiation therapy
   - Telecommunications optimization
4. On bridging the gap

Outline

1. Background
   - Nonlinear optimization
2. Methods
   - Interior methods
3. Applications
   - Optimization of radiation therapy
   - Telecommunications optimization
4. On bridging the gap

Outline

1. Background
   - Nonlinear optimization
2. Methods
   - Interior methods
3. Applications
   - Optimization of radiation therapy
   - Telecommunications optimization
4. On bridging the gap

Outline

1. Background
   - Nonlinear optimization
2. Methods
   - Interior methods
3. Applications
   - Optimization of radiation therapy
   - Telecommunications optimization
4. On bridging the gap

Outline

1. Background
   - Nonlinear optimization
2. Methods
   - Interior methods
3. Applications
   - Optimization of radiation therapy
   - Telecommunications optimization
4. On bridging the gap

Nonlinear optimization

A nonlinear optimization problem takes the form

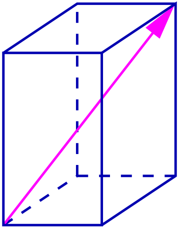
$$\begin{array}{lll} \underset{x \in \mathbb{R}^n}{\text{minimize}} & f(x) & \\ \text{subject to} & c_i(x) \geq 0, & i \in \mathcal{I}, \\ & c_i(x) = 0, & i \in \mathcal{E}, \end{array} \qquad \begin{array}{l} \mathcal{I} \bigcup \mathcal{E} = \{1, \dots, m\}, \\ \mathcal{I} \bigcap \mathcal{E} = \emptyset. \end{array}$$

where f and c_i, $i = 1, \dots, m$, are nonlinear smooth functions from $\mathbb{R}^n$ to $\mathbb{R}$.

The feasible region is denoted by F. In our case

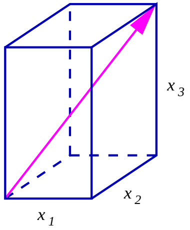
$$F = \{x \in \mathbb{R}^n : c_i(x) \geq 0,\ i \in \mathcal{I},\ c_i(x) = 0,\ i \in \mathcal{E}\}.$$

Example problem



Construct a box of volume 1 m^3 so that the space diagonal is minimized. What does it look like?

Formulation of example problem

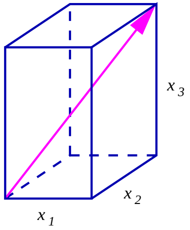


- Introduce variables x_i, $i = 1, \ldots 3$. We obtain

$$(P) \qquad \begin{array}{ll} \underset{x \in \mathbb{R}^3}{\text{minimize}} & x_1^2 + x_2^2 + x_3^2 \\ \text{subject to} & x_1 \cdot x_2 \cdot x_3 = 1, \\ & x_i \geq 0, \quad i = 1, 2, 3. \end{array}$$

- The problem is not convex.

Formulation of example problem



- Introduce variables x_i, $i = 1, \ldots 3$. We obtain

$$
(P) \quad \begin{array}{ll} \underset{x \in \mathbb{R}^3}{\text{minimize}} & x_1^2 + x_2^2 + x_3^2 \\ \text{subject to} & x_1 \cdot x_2 \cdot x_3 = 1, \\ & x_i \geq 0, \quad i = 1, 2, 3. \end{array}
$$

- The problem is not convex.

Alternative formulation of example problem

- We have the formulation

$$(P) \qquad \begin{array}{ll} \underset{x \in \mathbb{R}^3}{\text{minimize}} & x_1^2 + x_2^2 + x_3^2 \\ \text{subject to} & x_1 \cdot x_2 \cdot x_3 = 1, \\ & x_i \geq 0, \quad i = 1, 2, 3. \end{array}$$

- Replace $x_i \geq 0$, $i = 1, \dots, 3$ by $x_i > 0$, $i = 1, \dots, 3$.
- Let $y_i = \ln x_i$, $i = 1, 2, 3$, which gives

$$(P') \qquad \begin{array}{ll} \underset{y \in \mathbb{R}^3}{\text{minimize}} & e^{2y_1} + e^{2y_2} + e^{2y_3} \\ \text{subject to} & y_1 + y_2 + y_3 = 0. \end{array}$$

- This problem is convex.
- Is this a simpler problem to solve?

Alternative formulation of example problem

- We have the formulation

$$(P) \qquad \begin{array}{ll} \underset{x \in \mathbb{R}^3}{\text{minimize}} & x_1^2 + x_2^2 + x_3^2 \\ \text{subject to} & x_1 \cdot x_2 \cdot x_3 = 1, \\ & x_i \geq 0, \quad i = 1, 2, 3. \end{array}$$

- Replace $x_i \geq 0$, $i = 1, \ldots, 3$ by $x_i > 0$, $i = 1, \ldots, 3$.
- Let $y_i = \ln x_i$, $i = 1, 2, 3$, which gives

$$(P') \qquad \begin{array}{ll} \underset{y \in \mathbb{R}^3}{\text{minimize}} & e^{2y_1} + e^{2y_2} + e^{2y_3} \\ \text{subject to} & y_1 + y_2 + y_3 = 0. \end{array}$$

- This problem is convex.
- Is this a simpler problem to solve?

Alternative formulation of example problem

- We have the formulation

$$(P) \qquad \begin{array}{ll} \underset{x \in \mathbb{R}^3}{\text{minimize}} & x_1^2 + x_2^2 + x_3^2 \\ \text{subject to} & x_1 \cdot x_2 \cdot x_3 = 1, \\ & x_i \geq 0, \quad i = 1, 2, 3. \end{array}$$

- Replace $x_i \geq 0$, $i = 1, \dots, 3$ by $x_i > 0$, $i = 1, \dots, 3$.
- Let $y_i = \ln x_i$, $i = 1, 2, 3$, which gives

$$(P') \qquad \begin{array}{ll} \underset{y \in \mathbb{R}^3}{\text{minimize}} & e^{2y_1} + e^{2y_2} + e^{2y_3} \\ \text{subject to} & y_1 + y_2 + y_3 = 0. \end{array}$$

- This problem is convex.
- Is this a simpler problem to solve?

Alternative formulation of example problem

- We have the formulation

$$(P) \qquad \begin{array}{ll} \underset{x \in \mathbb{R}^3}{\text{minimize}} & x_1^2 + x_2^2 + x_3^2 \\ \text{subject to} & x_1 \cdot x_2 \cdot x_3 = 1, \\ & x_i \geq 0, \quad i = 1, 2, 3. \end{array}$$

- Replace $x_i \geq 0$, $i = 1, \ldots, 3$ by $x_i > 0$, $i = 1, \ldots, 3$.
- Let $y_i = \ln x_i$, $i = 1, 2, 3$, which gives

$$(P') \qquad \begin{array}{ll} \underset{y \in \mathbb{R}^3}{\text{minimize}} & e^{2y_1} + e^{2y_2} + e^{2y_3} \\ \text{subject to} & y_1 + y_2 + y_3 = 0. \end{array}$$

- This problem is convex.
- Is this a simpler problem to solve?

Alternative formulation of example problem

- We have the formulation

$$(P) \quad \begin{array}{ll} \underset{x\in\mathbb{R}^3}{\text{minimize}} & x_1^2 + x_2^2 + x_3^2 \\ \text{subject to} & x_1 \cdot x_2 \cdot x_3 = 1, \\ & x_i \geq 0, \quad i = 1,2,3. \end{array}$$

- Replace $x_i \geq 0$, $i = 1,\ldots,3$ by $x_i > 0$, $i = 1,\ldots,3$.
- Let $y_i = \ln x_i$, $i = 1,2,3$, which gives

$$(P') \quad \begin{array}{ll} \underset{y\in\mathbb{R}^3}{\text{minimize}} & e^{2y_1} + e^{2y_2} + e^{2y_3} \\ \text{subject to} & y_1 + y_2 + y_3 = 0. \end{array}$$

- This problem is convex.
- Is this a simpler problem to solve?

Applications of nonlinear optimization

- Nonlinear optimization arises in a wide range of areas.
- Particular areas of interest for this talk are:
  - Radiation therapy.
  - Telecommunications.
- The optimization problems are often very large.
- Problem structure is highly important.

Applications of nonlinear optimization

- Nonlinear optimization arises in a wide range of areas.
- Particular areas of interest for this talk are:
  - Radiation therapy.
  - Telecommunications.
- The optimization problems are often very large.
- Problem structure is highly important.

Applications of nonlinear optimization

- Nonlinear optimization arises in a wide range of areas.
- Particular areas of interest for this talk are:
  - Radiation therapy.
  - Telecommunications.
- The optimization problems are often very large.
- Problem structure is highly important.

Applications of nonlinear optimization

- Nonlinear optimization arises in a wide range of areas.
- Particular areas of interest for this talk are:
  - Radiation therapy.
  - Telecommunications.
- The optimization problems are often very large.
- Problem structure is highly important.

Applications of nonlinear optimization

- Nonlinear optimization arises in a wide range of areas.
- Particular areas of interest for this talk are:
  - Radiation therapy.
  - Telecommunications.
- The optimization problems are often very large.
- Problem structure is highly important.

Applications of nonlinear optimization

- Nonlinear optimization arises in a wide range of areas.
- Particular areas of interest for this talk are:
  - Radiation therapy.
  - Telecommunications.
- The optimization problems are often very large.
- Problem structure is highly important.

Outline

1. Background
   - Nonlinear optimization
2. Methods
   - Interior methods
3. Applications
   - Optimization of radiation therapy
   - Telecommunications optimization
4. On bridging the gap

Solution methods

- Solution methods are typically iterative methods that solve a sequence of simpler problems.
- Methods differ in terms of how complex subproblems that are formed.
- Many methods exist, e.g., interior methods, sequential quadratic programming methods etc.
- Rule of thumb: Second-derivatives are useful.

Solution methods

- Solution methods are typically iterative methods that solve a sequence of simpler problems.
- Methods differ in terms of how complex subproblems that are formed.
- Many methods exist, e.g., interior methods, sequential quadratic programming methods etc.
- Rule of thumb: Second-derivatives are useful.

Solution methods

- Solution methods are typically iterative methods that solve a sequence of simpler problems.
- Methods differ in terms of how complex subproblems that are formed.
- Many methods exist, e.g., interior methods, sequential quadratic programming methods etc.
- Rule of thumb: Second-derivatives are useful.

Solution methods

- Solution methods are typically iterative methods that solve a sequence of simpler problems.
- Methods differ in terms of how complex subproblems that are formed.
- Many methods exist, e.g., interior methods, sequential quadratic programming methods etc.
- Rule of thumb: Second-derivatives are useful.

Solution methods

- Sequential-quadratic programming (SQP) methods.
  - Local quadratic models of the problem are made.
  - Subproblem is a constrained quadratic program.
  - “Hard” prediction of active constraints.
  - Subproblem may be warmstarted.
- Interior methods.
  - Linearizations of perturbed optimality conditions are made.
  - Subproblem is a system of linear equations.
  - “Soft” prediction of active constraints.
  - Warm start is not easy.

Solution methods

- Sequential-quadratic programming (SQP) methods.
  - Local quadratic models of the problem are made.
  - Subproblem is a constrained quadratic program.
  - “Hard” prediction of active constraints.
  - Subproblem may be warmstarted.
- Interior methods.
  - Linearizations of perturbed optimality conditions are made.
  - Subproblem is a system of linear equations.
  - “Soft” prediction of active constraints.
  - Warm start is not easy.

Solution methods

- Sequential-quadratic programming (SQP) methods.
  - Local quadratic models of the problem are made.
  - Subproblem is a constrained quadratic program.
  - "Hard" prediction of active constraints.
  - Subproblem may be warmstarted.
- Interior methods.
  - Linearizations of perturbed optimality conditions are made.
  - Subproblem is a system of linear equations.
  - "Soft" prediction of active constraints.
  - Warm start is not easy.

Solution methods

- Sequential-quadratic programming (SQP) methods.
  - Local quadratic models of the problem are made.
  - Subproblem is a constrained quadratic program.
  - “Hard” prediction of active constraints.
  - Subproblem may be warmstarted.
- Interior methods.
  - Linearizations of perturbed optimality conditions are made.
  - Subproblem is a system of linear equations.
  - “Soft” prediction of active constraints.
  - Warm start is not easy.

Solution methods

- Sequential-quadratic programming (SQP) methods.
  - Local quadratic models of the problem are made.
  - Subproblem is a constrained quadratic program.
  - “Hard” prediction of active constraints.
  - Subproblem may be warmstarted.
- Interior methods.
  - Linearizations of perturbed optimality conditions are made.
  - Subproblem is a system of linear equations.
  - “Soft” prediction of active constraints.
  - Warm start is not easy.

Solution methods

- Sequential-quadratic programming (SQP) methods.
  - Local quadratic models of the problem are made.
  - Subproblem is a constrained quadratic program.
  - “Hard” prediction of active constraints.
  - Subproblem may be warmstarted.
- Interior methods.
  - Linearizations of perturbed optimality conditions are made.
  - Subproblem is a system of linear equations.
  - “Soft” prediction of active constraints.
  - Warm start is not easy.

Solution methods

- Sequential-quadratic programming (SQP) methods.
  - Local quadratic models of the problem are made.
  - Subproblem is a constrained quadratic program.
  - “Hard” prediction of active constraints.
  - Subproblem may be warmstarted.
- Interior methods.
  - Linearizations of perturbed optimality conditions are made.
  - Subproblem is a system of linear equations.
  - “Soft” prediction of active constraints.
  - Warm start is not easy.

Solution methods

- Sequential-quadratic programming (SQP) methods.
  - Local quadratic models of the problem are made.
  - Subproblem is a constrained quadratic program.
  - “Hard” prediction of active constraints.
  - Subproblem may be warmstarted.
- Interior methods.
  - Linearizations of perturbed optimality conditions are made.
  - Subproblem is a system of linear equations.
  - “Soft” prediction of active constraints.
  - Warm start is not easy.

Solution methods

- Sequential-quadratic programming (SQP) methods.
  - Local quadratic models of the problem are made.
  - Subproblem is a constrained quadratic program.
  - “Hard” prediction of active constraints.
  - Subproblem may be warmstarted.
- Interior methods.
  - Linearizations of perturbed optimality conditions are made.
  - Subproblem is a system of linear equations.
  - “Soft” prediction of active constraints.
  - Warm start is not easy.

Solution methods

- Sequential-quadratic programming (SQP) methods.
  - Local quadratic models of the problem are made.
  - Subproblem is a constrained quadratic program.
  - “Hard” prediction of active constraints.
  - Subproblem may be warmstarted.
- Interior methods.
  - Linearizations of perturbed optimality conditions are made.
  - Subproblem is a system of linear equations.
  - “Soft” prediction of active constraints.
  - Warm start is not easy.

Derivative information

- First-derivative methods are often not efficient enough.
- SQP methods and interior methods are second-derivative methods.
- An alternative to exact second derivatives are quasi-Newton methods.
- Stronger convergence properties for exact second derivatives.
- Exact second derivatives expected to be more efficient in practice.
- Exact second derivatives requires handling of nonconvexity.

Derivative information

- First-derivative methods are often not efficient enough.
- SQP methods and interior methods are second-derivative methods.
- An alternative to exact second derivatives are quasi-Newton methods.
- Stronger convergence properties for exact second derivatives.
- Exact second derivatives expected to be more efficient in practice.
- Exact second derivatives requires handling of nonconvexity.

Derivative information

- First-derivative methods are often not efficient enough.
- SQP methods and interior methods are second-derivative methods.
- An alternative to exact second derivatives are quasi-Newton methods.
- Stronger convergence properties for exact second derivatives.
- Exact second derivatives expected to be more efficient in practice.
- Exact second derivatives requires handling of nonconvexity.

Derivative information

- First-derivative methods are often not efficient enough.
- SQP methods and interior methods are second-derivative methods.
- An alternative to exact second derivatives are quasi-Newton methods.
- Stronger convergence properties for exact second derivatives.
- Exact second derivatives expected to be more efficient in practice.
- Exact second derivatives requires handling of nonconvexity.

Derivative information

- First-derivative methods are often not efficient enough.
- SQP methods and interior methods are second-derivative methods.
- An alternative to exact second derivatives are quasi-Newton methods.
- Stronger convergence properties for exact second derivatives.
- Exact second derivatives expected to be more efficient in practice.
- Exact second derivatives requires handling of nonconvexity.

Derivative information

- First-derivative methods are often not efficient enough.
- SQP methods and interior methods are second-derivative methods.
- An alternative to exact second derivatives are quasi-Newton methods.
- Stronger convergence properties for exact second derivatives.
- Exact second derivatives expected to be more efficient in practice.
- Exact second derivatives requires handling of nonconvexity.

Outline

1. Background
   - Nonlinear optimization
2. Methods
   - Interior methods
3. Applications
   - Optimization of radiation therapy
   - Telecommunications optimization
4. On bridging the gap

Inequality-constrained nonlinear optimization problem

For simplicity, consider the inequality constrained optimization problem

$$\begin{array}{ll} \underset{x \in \mathbb{R}^n}{\text{minimize}} & f(x) \\ \text{subject to} & c_i(x) \geq 0, \ i = 1, \ldots, m, \end{array}$$

where $f : \mathbb{R}^n \rightarrow \mathbb{R}$ and $c : \mathbb{R}^n \rightarrow \mathbb{R}^m$ are twice-continuously differentiable functions.

We will consider an interior method.

- The underlying idea of an interior method is a barrier transformation.
- In a logarithmic barrier transformation, the constraint $c_i(x) \geq 0$ is replaced by a barrier term $-\ln(c_i(x))$, which is added to the objective function.
- A positive barrier parameter μ determines the weight which is put on the barrier term.
- The resulting logarithmic barrier function is

$$B_\mu(x) = f(x) - \mu \sum_{i=1}^{m} \ln c_i(x).$$

Barrier transformation

- The underlying idea of an interior method is a barrier transformation.
- In a logarithmic barrier transformation, the constraint $c_i(x) \geq 0$ is replaced by a barrier term $-\ln(c_i(x))$, which is added to the objective function.
- A positive barrier parameter μ determines the weight which is put on the barrier term.
- The resulting logarithmic barrier function is

$$B_\mu(x) = f(x) - \mu \sum_{i=1}^{m} \ln c_i(x).$$

Barrier transformation

- The underlying idea of an interior method is a barrier transformation.
- In a logarithmic barrier transformation, the constraint $c_i(x) \geq 0$ is replaced by a barrier term $-\ln(c_i(x))$, which is added to the objective function.
- A positive barrier parameter μ determines the weight which is put on the barrier term.
- The resulting logarithmic barrier function is

$$B_\mu(x) = f(x) - \mu \sum_{i=1}^{m} \ln c_i(x).$$

Barrier transformation

- The underlying idea of an interior method is a barrier transformation.
- In a logarithmic barrier transformation, the constraint $c_i(x) \geq 0$ is replaced by a barrier term $-\ln(c_i(x))$, which is added to the objective function.
- A positive barrier parameter μ determines the weight which is put on the barrier term.
- The resulting logarithmic barrier function is

$$B_\mu(x) = f(x) - \mu \sum_{i=1}^{m} \ln c_i(x).$$

Geometric illustration of example problem

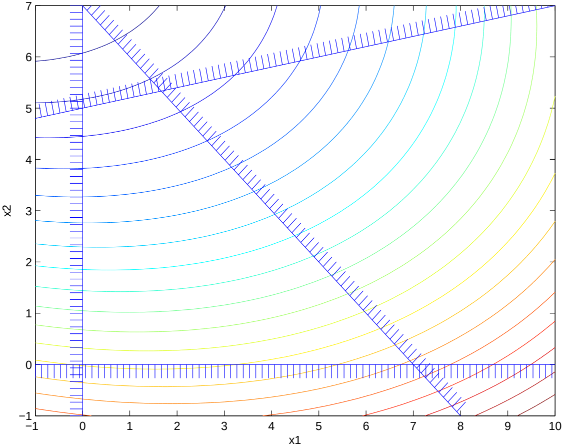


Illustration of primal barrier problem

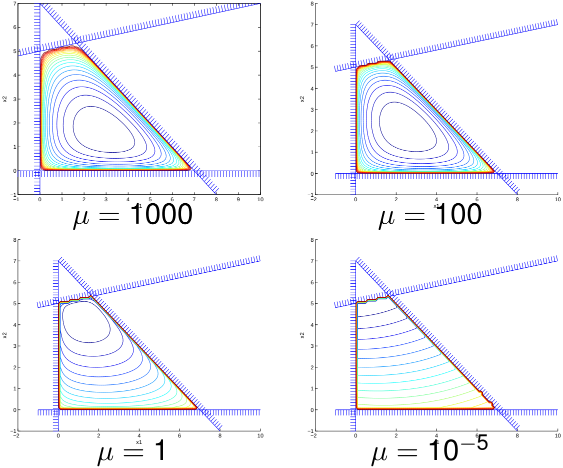
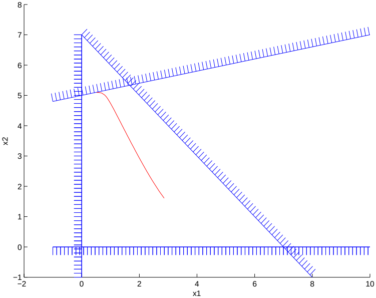
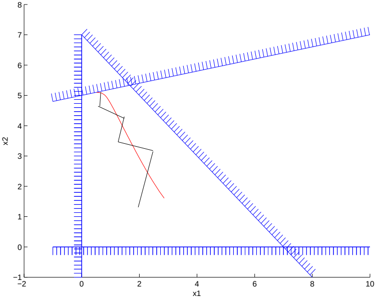


Illustration of primal part of barrier trajectory

An interior method approximately follows the barrier trajectory.



The trajectory



Generated iterates

Interior method iteration

- For a given $x \in \mathbb{R}^n$, $\lambda \in \mathbb{R}^m$ and $\mu > 0$, a primal-dual interior method solves the system

$$\begin{pmatrix} H(x,\lambda) & -A(x)^T \\ \Lambda A(x) & C(x) \end{pmatrix} \begin{pmatrix} \Delta x \\ \Delta\lambda \end{pmatrix} = - \begin{pmatrix} g(x) - A(x)^T\lambda \\ \Lambda C(x)e - \mu e \end{pmatrix},$$

  where $H(x,\lambda) = \nabla^2{}_{xx}(f(x) - \lambda^T c(x))$, $A(x)$ denotes the Jacobian of $c(x)$, e is the vector of ones, $C(x) = \text{diag}(c(x))$ and $\Lambda = \text{diag}(\lambda)$.
- Interior method in the sense that $c(x) > 0$ and $\lambda > 0$.
- Suppressing indices "x" and "λ", we obtain the augmented system

$$\begin{pmatrix} H & -A^T \\ \Lambda A & C \end{pmatrix} \begin{pmatrix} \Delta x \\ \Delta\lambda \end{pmatrix} = - \begin{pmatrix} g - A^T\lambda \\ \Lambda Ce - \mu e \end{pmatrix}.$$

Interior method iteration

- For a given $x \in \mathbb{R}^n$, $\lambda \in \mathbb{R}^m$ and $\mu > 0$, a primal-dual interior method solves the system

$$\begin{pmatrix} H(x,\lambda) & -A(x)^T \\ \Lambda A(x) & C(x) \end{pmatrix} \begin{pmatrix} \Delta x \\ \Delta \lambda \end{pmatrix} = - \begin{pmatrix} g(x) - A(x)^T \lambda \\ \Lambda C(x) e - \mu e \end{pmatrix},$$

  where $H(x,\lambda) = \nabla^2{}_{xx}(f(x) - \lambda^T c(x))$, $A(x)$ denotes the Jacobian of $c(x)$, e is the vector of ones, $C(x) = \mathrm{diag}(c(x))$ and $\Lambda = \mathrm{diag}(\lambda)$.
- Interior method in the sense that $c(x) > 0$ and $\lambda > 0$.
- Suppressing indices "x" and "λ", we obtain the augmented system

$$\begin{pmatrix} H & -A^T \\ \Lambda A & C \end{pmatrix} \begin{pmatrix} \Delta x \\ \Delta \lambda \end{pmatrix} = - \begin{pmatrix} g - A^T \lambda \\ \Lambda C e - \mu e \end{pmatrix}.$$

Interior method iteration

- For a given $x \in \mathbb{R}^n$, $\lambda \in \mathbb{R}^m$ and $\mu > 0$, a primal-dual interior method solves the system

$$\begin{pmatrix} H(x,\lambda) & -A(x)^T \\ \Lambda A(x) & C(x) \end{pmatrix} \begin{pmatrix} \Delta x \\ \Delta\lambda \end{pmatrix} = -\begin{pmatrix} g(x) - A(x)^T\lambda \\ \Lambda C(x)e - \mu e \end{pmatrix},$$

  where $H(x,\lambda) = \nabla^2{}_{xx}(f(x) - \lambda^T c(x))$, $A(x)$ denotes the Jacobian of $c(x)$, e is the vector of ones, $C(x) = \text{diag}(c(x))$ and $\Lambda = \text{diag}(\lambda)$.
- Interior method in the sense that $c(x) > 0$ and $\lambda > 0$.
- Suppressing indices "x" and "λ", we obtain the augmented system

$$\begin{pmatrix} H & -A^T \\ \Lambda A & C \end{pmatrix} \begin{pmatrix} \Delta x \\ \Delta\lambda \end{pmatrix} = -\begin{pmatrix} g - A^T\lambda \\ \Lambda Ce - \mu e \end{pmatrix}.$$

Outline

Outline

Radiation therapy

- Treatment of cancer is a very important task.
- Radiation therapy is one of the most powerful methods of treatment. In Sweden 30% of the cancer patients are treated with radiation therapy.
- The radiation may be optimized to improve performance of radiation.
- Hence, behind this important medical application is an optimization problem.

Radiation therapy

- Treatment of cancer is a very important task.
- Radiation therapy is one of the most powerful methods of treatment. In Sweden 30% of the cancer patients are treated with radiation therapy.
- The radiation may be optimized to improve performance of radiation.
- Hence, behind this important medical application is an optimization problem.

Radiation therapy

- Treatment of cancer is a very important task.
- Radiation therapy is one of the most powerful methods of treatment. In Sweden 30% of the cancer patients are treated with radiation therapy.
- The radiation may be optimized to improve performance of radiation.
- Hence, behind this important medical application is an optimization problem.

Radiation therapy

- Treatment of cancer is a very important task.
- Radiation therapy is one of the most powerful methods of treatment. In Sweden 30% of the cancer patients are treated with radiation therapy.
- The radiation may be optimized to improve performance of radiation.
- Hence, behind this important medical application is an optimization problem.

Radiation therapy

Treatment of cancer by radiation therapy means that the patient is subject to radiation by a particle beam.

The main parts of a treatment unit are:

- the particle accelerator, which creates the beam;
- beam optical components, which direct the beam;
- the supporting gantry, which gives the beam angle of incidence;
- the treatment head, which modulates the intensity of the beam.

Radiation therapy

Treatment of cancer by radiation therapy means that the patient is subject to radiation by a particle beam.

The main parts of a treatment unit are:

- the particle accelerator, which creates the beam;
- beam optical components, which direct the beam;
- the supporting gantry, which gives the beam angle of incidence;
- the treatment head, which modulates the intensity of the beam.

Radiation therapy

Treatment of cancer by radiation therapy means that the patient is subject to radiation by a particle beam.

The main parts of a treatment unit are:

- the particle accelerator, which creates the beam;
- beam optical components, which direct the beam;
- the supporting gantry, which gives the beam angle of incidence;
- the treatment head, which modulates the intensity of the beam.

Radiation therapy

Treatment of cancer by radiation therapy means that the patient is subject to radiation by a particle beam.

The main parts of a treatment unit are:

- the particle accelerator, which creates the beam;
- beam optical components, which direct the beam;
- the supporting gantry, which gives the beam angle of incidence;
- the treatment head, which modulates the intensity of the beam.

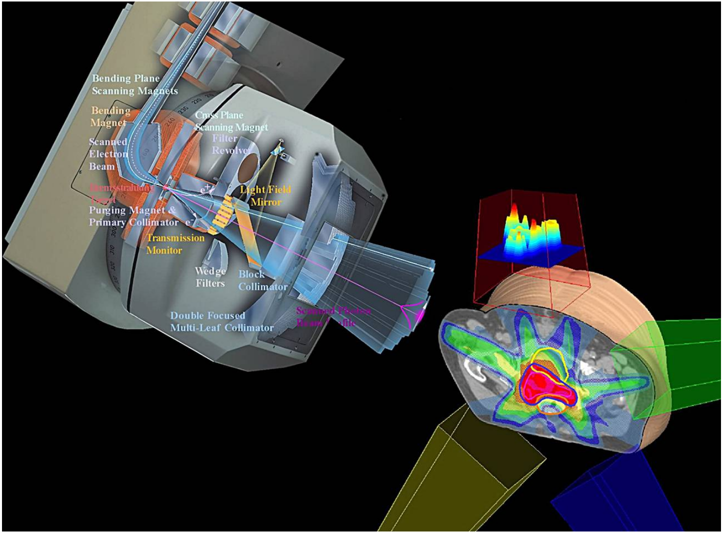
Radiation therapy

Treatment of cancer by radiation therapy means that the patient is subject to radiation by a particle beam.

The main parts of a treatment unit are:

- the particle accelerator, which creates the beam;
- beam optical components, which direct the beam;
- the supporting gantry, which gives the beam angle of incidence;
- the treatment head, which modulates the intensity of the beam.

Radiation treatment unit



Aim of radiation

- The aim of the radiation is typically to give a treatment that leads to a desirable dose distribution in the patient.
- Typically, high dose is desired in the tumor cells, and low dose in the other cells.
- In particular, certain organs are very sensitive to radiation and must have a low dose level, e.g., the spine.
- Hence, a desired dose distribution can be specified, and the question is how to achieve this distribution.
- This is an inverse problem in that the desired result of the radiation is known, but the treatment plan has to be designed.

Aim of radiation

- The aim of the radiation is typically to give a treatment that leads to a desirable dose distribution in the patient.
- Typically, high dose is desired in the tumor cells, and low dose in the other cells.
- In particular, certain organs are very sensitive to radiation and must have a low dose level, e.g., the spine.
- Hence, a desired dose distribution can be specified, and the question is how to achieve this distribution.
- This is an inverse problem in that the desired result of the radiation is known, but the treatment plan has to be designed.

Aim of radiation

- The aim of the radiation is typically to give a treatment that leads to a desirable dose distribution in the patient.
- Typically, high dose is desired in the tumor cells, and low dose in the other cells.
- In particular, certain organs are very sensitive to radiation and must have a low dose level, e.g., the spine.
- Hence, a desired dose distribution can be specified, and the question is how to achieve this distribution.
- This is an inverse problem in that the desired result of the radiation is known, but the treatment plan has to be designed.

Aim of radiation

- The aim of the radiation is typically to give a treatment that leads to a desirable dose distribution in the patient.
- Typically, high dose is desired in the tumor cells, and low dose in the other cells.
- In particular, certain organs are very sensitive to radiation and must have a low dose level, e.g., the spine.
- Hence, a desired dose distribution can be specified, and the question is how to achieve this distribution.
- This is an inverse problem in that the desired result of the radiation is known, but the treatment plan has to be designed.

Aim of radiation

- The aim of the radiation is typically to give a treatment that leads to a desirable dose distribution in the patient.
- Typically, high dose is desired in the tumor cells, and low dose in the other cells.
- In particular, certain organs are very sensitive to radiation and must have a low dose level, e.g., the spine.
- Hence, a desired dose distribution can be specified, and the question is how to achieve this distribution.
- This is an inverse problem in that the desired result of the radiation is known, but the treatment plan has to be designed.

Formulation of optimization problem

- A radiation treatment is typically given as a series of radiations.
- For an individual treatment, the performance depends on
  - the beam angle of incidence, which is governed by the supporting gantry; and
  - the intensity modulation of the beam, which is governed by the treatment head.
- One may now formulate an optimization problem, where the variables are the beam angles of incidence and the intensity modulations of the beams.
- In this talk, we assume that the beam angles of incidence are fixed.

Formulation of optimization problem

- A radiation treatment is typically given as a series of radiations.
- For an individual treatment, the performance depends on
  - the beam angle of incidence, which is governed by the supporting gantry; and
  - the intensity modulation of the beam, which is governed by the treatment head.
- One may now formulate an optimization problem, where the variables are the beam angles of incidence and the intensity modulations of the beams.
- In this talk, we assume that the beam angles of incidence are fixed.

Formulation of optimization problem

- A radiation treatment is typically given as a series of radiations.
- For an individual treatment, the performance depends on
  - the beam angle of incidence, which is governed by the supporting gantry; and
  - the intensity modulation of the beam, which is governed by the treatment head.
- One may now formulate an optimization problem, where the variables are the beam angles of incidence and the intensity modulations of the beams.
- In this talk, we assume that the beam angles of incidence are fixed.

Formulation of optimization problem

- A radiation treatment is typically given as a series of radiations.
- For an individual treatment, the performance depends on
  - the beam angle of incidence, which is governed by the supporting gantry; and
  - the intensity modulation of the beam, which is governed by the treatment head.
- One may now formulate an optimization problem, where the variables are the beam angles of incidence and the intensity modulations of the beams.
- In this talk, we assume that the beam angles of incidence are fixed.

Formulation of optimization problem

- A radiation treatment is typically given as a series of radiations.
- For an individual treatment, the performance depends on
  - the beam angle of incidence, which is governed by the supporting gantry; and
  - the intensity modulation of the beam, which is governed by the treatment head.
- One may now formulate an optimization problem, where the variables are the beam angles of incidence and the intensity modulations of the beams.
- In this talk, we assume that the beam angles of incidence are fixed.

Formulation of optimization problem

- A radiation treatment is typically given as a series of radiations.
- For an individual treatment, the performance depends on
  - the beam angle of incidence, which is governed by the supporting gantry; and
  - the intensity modulation of the beam, which is governed by the treatment head.
- One may now formulate an optimization problem, where the variables are the beam angles of incidence and the intensity modulations of the beams.
- In this talk, we assume that the beam angles of incidence are fixed.

Characteristics of the optimization problem

- The resulting optimization problem is a large-scale nonlinear optimization problem, typically with a large number of degrees of freedom at the solution.
- Many different objective functions have been proposed. The above problem characteristics hold.

Characteristics of the optimization problem

- The resulting optimization problem is a large-scale nonlinear optimization problem, typically with a large number of degrees of freedom at the solution.
- Many different objective functions have been proposed. The above problem characteristics hold.

Optimization of radiation therapy

Joint research project between

KTH and RaySearch Laboratories AB.

Financially supported by the Swedish Research Council.

Industrial graduate student: Fredrik Carlsson. (PhD April 2008)

Solution method

A simplified bound-constrained problem may be posed as

$$\begin{array}{ll} \underset{x \in \mathbb{R}^n}{\text{minimize}} & f(x) \\ \text{subject to} & l \leq x \leq u. \end{array}$$

- Large-scale problem solved in few (˜20) iterations using a quasi-Newton SQP method.
- Difficulty: "Jagged" solutions for more accurate plans.
- Idea: Use second-derivatives and an interior method to obtain fast convergence and smooth solutions.
    - Good news: Faster convergence.
    - Bad news: Increased jaggedness.
- Not following the folklore.

Solution method

A simplified bound-constrained problem may be posed as

$$\begin{array}{ll} \underset{x \in \mathbb{R}^n}{\text{minimize}} & f(x) \\ \text{subject to} & l \le x \le u. \end{array}$$

- Large-scale problem solved in few (˜20) iterations using a quasi-Newton SQP method.
- Difficulty: “Jagged” solutions for more accurate plans.
- Idea: Use second-derivatives and an interior method to obtain fast convergence and smooth solutions.
  - Good news: Faster convergence.
  - Bad news: Increased jaggedness.
- Not following the folklore.

Solution method

A simplified bound-constrained problem may be posed as

$$\begin{array}{ll} \underset{x \in \mathbb{R}^n}{\text{minimize}} & f(x) \\ \text{subject to} & l \leq x \leq u. \end{array}$$

- Large-scale problem solved in few (˜20) iterations using a quasi-Newton SQP method.
- Difficulty: “Jagged” solutions for more accurate plans.
- Idea: Use second-derivatives and an interior method to obtain fast convergence and smooth solutions.
  - Good news: Faster convergence.
  - Bad news: Increased jaggedness.
- Not following the folklore.

Solution method

A simplified bound-constrained problem may be posed as

$$\begin{array}{ll} \underset{x \in \mathbb{R}^n}{\text{minimize}} & f(x) \\ \text{subject to} & l \leq x \leq u. \end{array}$$

- Large-scale problem solved in few (˜20) iterations using a quasi-Newton SQP method.
- Difficulty: “Jagged” solutions for more accurate plans.
- Idea: Use second-derivatives and an interior method to obtain fast convergence and smooth solutions.
  - Good news: Faster convergence.
  - Bad news: Increased jaggedness.
- Not following the folklore.

Solution method

A simplified bound-constrained problem may be posed as

$$\begin{array}{ll} \underset{x \in \mathbb{R}^n}{\text{minimize}} & f(x) \\ \text{subject to} & l \leq x \leq u. \end{array}$$

- Large-scale problem solved in few (˜20) iterations using a quasi-Newton SQP method.
- Difficulty: “Jagged” solutions for more accurate plans.
- Idea: Use second-derivatives and an interior method to obtain fast convergence and smooth solutions.
  - Good news: Faster convergence.
  - Bad news: Increased jaggedness.
- Not following the folklore.

Solution method

A simplified bound-constrained problem may be posed as

$$\begin{array}{ll} \underset{x \in \mathbb{R}^n}{\text{minimize}} & f(x) \\ \text{subject to} & l \le x \le u. \end{array}$$

- Large-scale problem solved in few (˜20) iterations using a quasi-Newton SQP method.
- Difficulty: "Jagged" solutions for more accurate plans.
- Idea: Use second-derivatives and an interior method to obtain fast convergence and smooth solutions.
  - Good news: Faster convergence.
  - Bad news: Increased jaggedness.
- Not following the folklore.

Radiation therapy and the conjugate-gradient method

- Why does a quasi-Newton sequential quadratic programming method do so well on these problems?
- The answer lies in the problem structure.
- Simplify further, consider a quadratic approximation of the objective function and eliminate the constraints.

$$\underset{x \in \mathbb{R}^n}{\text{minimize}} \quad \tfrac{1}{2} x^T H x + c^T x$$

  where $H = H^T \succeq 0$.
- Quasi-Newton methods and the conjugate-gradient method are equivalent on this problem.
- The conjugate-gradient method minimizes in directions corresponding to large eigenvalues first.

Radiation therapy and the conjugate-gradient method

- Why does a quasi-Newton sequential quadratic programming method do so well on these problems?
- The answer lies in the problem structure.
- Simplify further, consider a quadratic approximation of the objective function and eliminate the constraints.

$$\underset{x \in \mathbb{R}^n}{\text{minimize}} \quad \tfrac{1}{2} x^T H x + c^T x$$

where $H = H^T \succeq 0$.

- Quasi-Newton methods and the conjugate-gradient method are equivalent on this problem.
- The conjugate-gradient method minimizes in directions corresponding to large eigenvalues first.

Radiation therapy and the conjugate-gradient method

- Why does a quasi-Newton sequential quadratic programming method do so well on these problems?
- The answer lies in the problem structure.
- Simplify further, consider a quadratic approximation of the objective function and eliminate the constraints.

$$\underset{x \in \mathbb{R}^n}{\text{minimize}} \qquad \tfrac{1}{2} x^T H x + c^T x$$

  where $H = H^T \succeq 0$.
- Quasi-Newton methods and the conjugate-gradient method are equivalent on this problem.
- The conjugate-gradient method minimizes in directions corresponding to large eigenvalues first.

Radiation therapy and the conjugate-gradient method

- Why does a quasi-Newton sequential quadratic programming method do so well on these problems?
- The answer lies in the problem structure.
- Simplify further, consider a quadratic approximation of the objective function and eliminate the constraints.

$$\underset{x \in \mathbb{R}^n}{\text{minimize}} \qquad \tfrac{1}{2} x^T H x + c^T x$$

  where $H = H^T \succeq 0$.
- Quasi-Newton methods and the conjugate-gradient method are equivalent on this problem.
- The conjugate-gradient method minimizes in directions corresponding to large eigenvalues first.

Radiation therapy and the conjugate-gradient method

- Why does a quasi-Newton sequential quadratic programming method do so well on these problems?
- The answer lies in the problem structure.
- Simplify further, consider a quadratic approximation of the objective function and eliminate the constraints.

$$\underset{x \in \mathbb{R}^n}{\text{minimize}} \qquad \tfrac{1}{2} x^T H x + c^T x$$

  where $H = H^T \succeq 0$.
- Quasi-Newton methods and the conjugate-gradient method are equivalent on this problem.
- The conjugate-gradient method minimizes in directions corresponding to large eigenvalues first.

Radiation therapy and the conjugate-gradient method

- The conjugate-gradient method minimizes in directions corresponding to large eigenvalues first.
- Our simplified problem has few large eigenvalues, corresponding to smooth solutions.
- Many small eigenvalues that correspond to jagged solutions.
- The conjugate-gradient method takes a desirable path to the solution.
- Additional properties of the solution, not seen in the formulation, are important.

Radiation therapy and the conjugate-gradient method

- The conjugate-gradient method minimizes in directions corresponding to large eigenvalues first.
- Our simplified problem has few large eigenvalues, corresponding to smooth solutions.
- Many small eigenvalues that correspond to jagged solutions.
- The conjugate-gradient method takes a desirable path to the solution.
- Additional properties of the solution, not seen in the formulation, are important.

Radiation therapy and the conjugate-gradient method

- The conjugate-gradient method minimizes in directions corresponding to large eigenvalues first.
- Our simplified problem has few large eigenvalues, corresponding to smooth solutions.
- Many small eigenvalues that correspond to jagged solutions.
- The conjugate-gradient method takes a desirable path to the solution.
- Additional properties of the solution, not seen in the formulation, are important.

Radiation therapy and the conjugate-gradient method

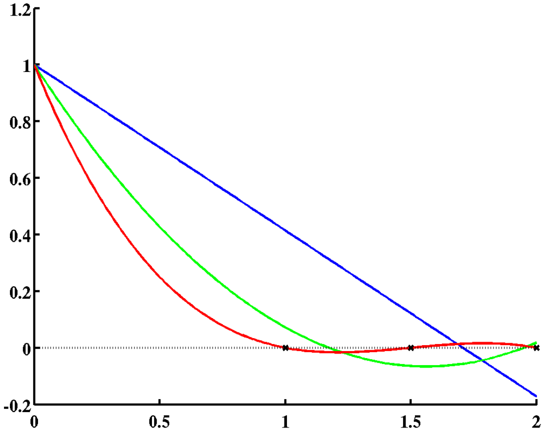
- The conjugate-gradient method minimizes in directions corresponding to large eigenvalues first.
- Our simplified problem has few large eigenvalues, corresponding to smooth solutions.
- Many small eigenvalues that correspond to jagged solutions.
- The conjugate-gradient method takes a desirable path to the solution.
- Additional properties of the solution, not seen in the formulation, are important.

Radiation therapy and the conjugate-gradient method

- The conjugate-gradient method minimizes in directions corresponding to large eigenvalues first.
- Our simplified problem has few large eigenvalues, corresponding to smooth solutions.
- Many small eigenvalues that correspond to jagged solutions.
- The conjugate-gradient method takes a desirable path to the solution.
- Additional properties of the solution, not seen in the formulation, are important.

Polynomials for well-conditioned example problem

Example problem with $\lambda = (2, 1.5, 1)^T$ and $\xi^{(0)} = (1, 1, 1)^T$, with $\lambda = \text{eig}(H)$, $\xi = x - x^*$.



Behavior of the conjugate gradient subproblems

$$\begin{array}{ll} \underset{\xi\in\mathbb{R}^n,\zeta\in\mathbb{R}^k}{\text{minimize}} & \frac{1}{2}\sum_{i=1}^{n}\lambda_i\xi_i^2 \\ \text{subject to} & \xi_i = \prod_{l=1}^{k}\left(1-\frac{\lambda_i}{\zeta_l}\right)\xi_i^{(0)}, \quad i=1,\dots,n, \end{array}$$

The optimal solution $\xi^{(k)}$ will tend to have smaller components $\xi_i^{(k)}$ for i such that λ_i is large and/or $\xi_i^{(0)}$ is large.

Nonlinear dependency of $\xi^{(k)}$ on λ and $\xi^{(0)}$.

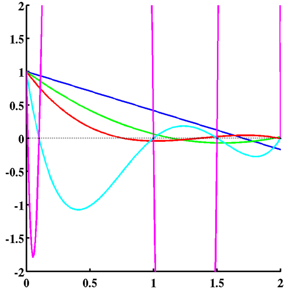
We are interested in the ill-conditioned case, when H has relatively few large eigenvalues.

Empirical behavior on ill-conditioned problem

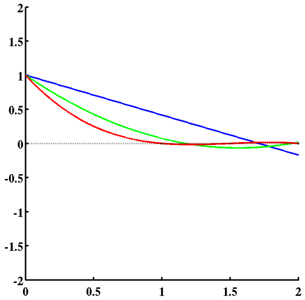
Based on the ill-conditioned example, we make the following observations:

- The conjugate-gradient method initially “approximately” solves the problem given by the large eigenvalues only.
- During this initial stage, the method “almost” behaves as the conjugate method applied to the smaller problem given by the large eigenvalues only.
- During this initial stage, the residuals associated with small eigenvalues of H are “almost” unchanged.
- The ill-conditioning of the problem does not appear in this initial stage.

Polynomials for ill-conditioned example problem



Polynomials for problem with $\lambda = (2, 1.5, 1, 0.1, 0.01)^T$ and $\xi^{(0)} = (1, 1, 1, 1, 1)^T$.



Polynomials for problem with $\lambda = (2, 1.5, 1)^T$ and $\xi^{(0)} = (1, 1, 1)^T$.

Empirical behavior on ill-conditioned problem

Based on the ill-conditioned example, we make the following observations:

- The conjugate-gradient method initially "approximately" solves the problem given by the large eigenvalues only.
- During this initial stage, the method "almost" behaves as the conjugate method applied to the smaller problem given by the large eigenvalues only.
- During this initial stage, the residuals associated with small eigenvalues of H are "almost" unchanged.
- The ill-conditioning of the problem does not appear in this initial stage.

Empirical behavior on ill-conditioned problem

Based on the ill-conditioned example, we make the following observations:

- The conjugate-gradient method initially "approximately" solves the problem given by the large eigenvalues only.
- During this initial stage, the method "almost" behaves as the conjugate method applied to the smaller problem given by the large eigenvalues only.
- During this initial stage, the residuals associated with small eigenvalues of H are "almost" unchanged.
- The ill-conditioning of the problem does not appear in this initial stage.

Empirical behavior on ill-conditioned problem

Based on the ill-conditioned example, we make the following observations:

- The conjugate-gradient method initially “approximately” solves the problem given by the large eigenvalues only.
- During this initial stage, the method “almost” behaves as the conjugate method applied to the smaller problem given by the large eigenvalues only.
- During this initial stage, the residuals associated with small eigenvalues of H are “almost” unchanged.
- The ill-conditioning of the problem does not appear in this initial stage.

Empirical behavior on ill-conditioned problem

Based on the ill-conditioned example, we make the following observations:

- The conjugate-gradient method initially “approximately” solves the problem given by the large eigenvalues only.
- During this initial stage, the method “almost” behaves as the conjugate method applied to the smaller problem given by the large eigenvalues only.
- During this initial stage, the residuals associated with small eigenvalues of H are “almost” unchanged.
- The ill-conditioning of the problem does not appear in this initial stage.

Optimization of radiation therapy

- Problem characteristics explained behavior of quasi-Newton method.
- Quasi-Newton method “better” than second-derivative method.
- Later reserach: Optimization directly on machine settings.
- Column generation used, mixed with machine parameter optimization.

Optimization of radiation therapy

- Problem characteristics explained behavior of quasi-Newton method.
- Quasi-Newton method “better” than second-derivative method.
- Later reserach: Optimization directly on machine settings.
- Column generation used, mixed with machine parameter optimization.

Optimization of radiation therapy

- Problem characteristics explained behavior of quasi-Newton method.
- Quasi-Newton method “better” than second-derivative method.
- Later reserach: Optimization directly on machine settings.
- Column generation used, mixed with machine parameter optimization.

Optimization of radiation therapy

- Problem characteristics explained behavior of quasi-Newton method.
- Quasi-Newton method “better” than second-derivative method.
- Later reserach: Optimization directly on machine settings.
- Column generation used, mixed with machine parameter optimization.

Outline

Optimization approaches to distributed multi-cell radio resource management

Research project within the

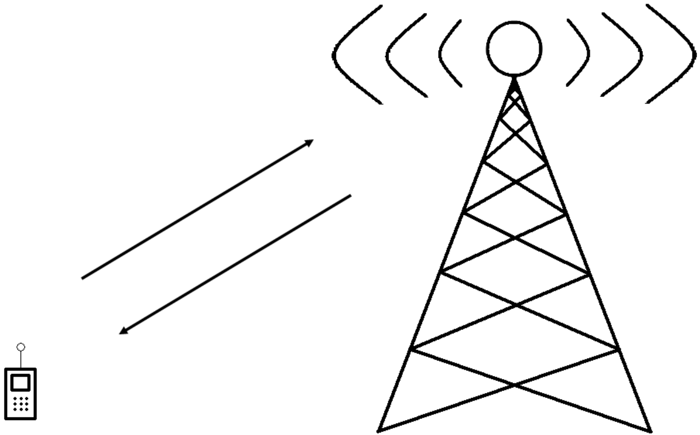
KTH Center for Industrial and Applied Mathematics (CIAM).

Industrial partner: Ericsson.

Financially supported by the
Swedish Foundation for Strategic Research.

Graduate student: Mikael Fallgren.

Radio resource management



Optimization problem

- Maximize throughput.
- Nonconvex problem.
- Convexification possible.
- Leads to loss of separability.

Question: How is this problem best solved?

Research in progress.

Optimization problem

- Maximize throughput.
- Nonconvex problem.
- Convexification possible.
- Leads to loss of separability.

Question: How is this problem best solved?

Research in progress.

Optimization problem

- Maximize throughput.
- Nonconvex problem.
- Convexification possible.
- Leads to loss of separability.

Question: How is this problem best solved?

Research in progress.

Optimization problem

- Maximize throughput.
- Nonconvex problem.
- Convexification possible.
- Leads to loss of separability.

Question: How is this problem best solved?

Research in progress.

Optimization problem

- Maximize throughput.
- Nonconvex problem.
- Convexification possible.
- Leads to loss of separability.

Question: How is this problem best solved?

Research in progress.

Optimization problem

- Maximize throughput.
- Nonconvex problem.
- Convexification possible.
- Leads to loss of separability.

Question: How is this problem best solved?

Research in progress.

Optimization problem

- Maximize throughput.
- Nonconvex problem.
- Convexification possible.
- Leads to loss of separability.

Question: How is this problem best solved?

Research in progress.

Outline

1. Background
   - Nonlinear optimization
2. Methods
   - Interior methods
3. Applications
   - Optimization of radiation therapy
   - Telecommunications optimization
4. On bridging the gap

Summary

A personal view on nonlinear optimization.

- Methods are very important.
- Applications give new challenges.
- Often two-way communication between method and application.
- Collaboration with application experts extremely important.

Summary

A personal view on nonlinear optimization.

- Methods are very important.
- Applications give new challenges.
- Often two-way communication between method and application.
- Collaboration with application experts extremely important.

Summary

A personal view on nonlinear optimization.

- Methods are very important.
- Applications give new challenges.
- Often two-way communication between method and application.
- Collaboration with application experts extremely important.

Summary

A personal view on nonlinear optimization.

- Methods are very important.
- Applications give new challenges.
- Often two-way communication between method and application.
- Collaboration with application experts extremely important.

On bridging the gap at KTH

Center for Industrial and Applied Mathematics (CIAM) at KTH.

- Research projects in mathematics with an applied profile.
- Joint projects with industry or other academic partners.
- "Umbrella" for a wide range of applications.
- Financially supported by the Swedish Foundation for Strategic Research.
- Interaction and collaboration is the key to success.
- One important goal is to bridge the gap between academia and industry.

On bridging the gap at KTH

Center for Industrial and Applied Mathematics (CIAM) at KTH.

- Research projects in mathematics with an applied profile.
- Joint projects with industry or other academic partners.
- "Umbrella" for a wide range of applications.
- Financially supported by the Swedish Foundation for Strategic Research.
- Interaction and collaboration is the key to success.
- One important goal is to bridge the gap between academia and industry.

On bridging the gap at KTH

Center for Industrial and Applied Mathematics (CIAM) at KTH.

- Research projects in mathematics with an applied profile.
- Joint projects with industry or other academic partners.
- “Umbrella” for a wide range of applications.
- Financially supported by the Swedish Foundation for Strategic Research.
- Interaction and collaboration is the key to success.
- One important goal is to bridge the gap between academia and industry.

On bridging the gap at KTH

Center for Industrial and Applied Mathematics (CIAM) at KTH.

- Research projects in mathematics with an applied profile.
- Joint projects with industry or other academic partners.
- “Umbrella” for a wide range of applications.
- Financially supported by the Swedish Foundation for Strategic Research.
- Interaction and collaboration is the key to success.
- One important goal is to bridge the gap between academia and industry.

On bridging the gap at KTH

Center for Industrial and Applied Mathematics (CIAM) at KTH.

- Research projects in mathematics with an applied profile.
- Joint projects with industry or other academic partners.
- “Umbrella” for a wide range of applications.
- Financially supported by the Swedish Foundation for Strategic Research.
- Interaction and collaboration is the key to success.
- One important goal is to bridge the gap between academia and industry.

On bridging the gap at KTH

Center for Industrial and Applied Mathematics (CIAM) at KTH.

- Research projects in mathematics with an applied profile.
- Joint projects with industry or other academic partners.
- “Umbrella” for a wide range of applications.
- Financially supported by the Swedish Foundation for Strategic Research.
- Interaction and collaboration is the key to success.
- One important goal is to bridge the gap between academia and industry.

The end

The presentation is based on joint research with various collaborators:

- Fredrik Carlsson
- Mikael Fallgren
- Philip Gill
- Josh Griffin
- ...

References: http://www.math.kth.se/~andersf

Thank you for your attention!