# ISEN 629: Engineering Optimization

## Lecture 26

Sergiy Butenko

Industrial and Systems Engineering
Texas A&M University

Fall 2007



## Duality in nonlinear optimization

Let $f$, $g$, $F$ be real–valued functions defined on $X \subseteq \mathbb{R}^m$, $Y \subseteq \mathbb{R}^n$ and $X \times Y \subseteq \mathbb{R}^m \times \mathbb{R}^n$, respectively.

Assume that the global minima and maxima which we address below do exist in all cases.

Suppose that $f(x) \leq g(y)$ for all $(x, y) \in X \times Y$. Then

$$\max_{x \in X} f(x) \leq \min_{y \in Y} g(y). \quad \text{(weak duality)}$$

Under certain conditions this inequality can be satisfied as an equality

$$\max_{x \in X} f(x) = \min_{y \in Y} g(y). \quad \text{(strong duality)}$$



## Duality in nonlinear optimization

It is easy to show that the following inequality holds:

$$\max_{y \in Y} \min_{x \in X} F(x, y) \leq \min_{x \in X} \max_{y \in Y} F(x, y).$$

Under certain conditions we can prove that

$$\max_{y \in Y} \min_{x \in X} F(x, y) = \min_{x \in X} \max_{y \in Y} F(x, y). \quad \text{(minimax)}$$



## Duality in nonlinear optimization

The point $(x^*, y^*) \in X \times Y$ is a saddlepoint of $F$ (with respect to maximizing in $Y$ and minimizing in $X$) if

$F(x^*, y) \leq F(x^*, y^*) \leq F(x, y^*)$ for every $(x, y) \in X \times Y$.

**Theorem**
*The point $(x^*, y^*) \in X \times Y$ is a saddlepoint of $F$ iff*

$$F(x^*, y^*) = \max_{y \in Y} \min_{x \in X} F(x, y) = \min_{x \in X} \max_{y \in Y} F(x, y).$$

**Proof.**
Let $(x^*, y^*)$ be a saddlepoint of $F$.

$$\begin{aligned} F(x^*, y^*) &= \min_{x \in X} F(x, y^*) \leq \max_{y \in Y} \min_{x \in X} F(x, y) \\ &\leq \min_{x \in X} \max_{y \in Y} F(x, y) \leq \max_{y \in Y} F(x^*, y) \\ &= F(x^*, y^*). \end{aligned}$$

$\square$

## Duality in nonlinear optimization

Consider the primal optimization problem

$$\begin{array}{ll} \min & f(x) \\ \text{s.t.} & g(x) \leq 0, \\ & h(x) = 0, \\ & x \in X, \end{array} \qquad \text{(P)}$$

where $g : \mathbb{R}^n \to \mathbb{R}^m,\ h : \mathbb{R}^n \to \mathbb{R}^p$ and $X \subseteq \mathbb{R}^n$.

The Lagrangian of (P) is

$$L(x, \lambda, \nu) = f(x) + \lambda^T g(x) + \nu^T h(x),$$

where

$$\lambda \in \mathbb{R}^m_+, \nu \in \mathbb{R}^p.$$



## Duality in nonlinear optimization

Note that for the Lagrangian

$$L(x, \lambda, \nu) = f(x) + \lambda^T g(x) + \nu^T h(x),$$

we have

$$\sup_{\lambda \geq 0, \nu} \{L(x, \lambda, \nu)\} = \begin{cases} f(x), & \text{if } g(x) \leq 0 \text{ and } h(x) = 0 \\ +\infty, & \text{otherwise,} \end{cases}$$

and the problem (P) can be restated in the form

$$\min_{x \in X} \max_{\lambda \geq 0, \nu} L(x, \lambda, \nu).$$



## Duality in nonlinear optimization

For $\lambda \geq 0$ and $\nu$ define the function

$$d(\lambda, \nu) = \min_{x \in X} L(x, \lambda, \nu)$$

which is concave (independently of the convexity of $f$).

Then the dual problem of (P) is defined by the following optimization problem:

$$\max_{\lambda \geq 0, \nu} d(\lambda, \nu) = \max_{\lambda \geq 0, \nu} \min_{x \in X} L(x, \lambda, \nu). \qquad \text{(D)}$$

The objective function $d(\lambda, \nu)$ of (D) is called the dual function.



## Duality in nonlinear optimization

### Theorem (Weak Duality Theorem)

*Let $x^*$ be a global minimum point of the primal problem. Then for every $\lambda \geq 0, \nu$*

$$d(\lambda, \nu) \leq d(\lambda^*, \nu^*) \leq f(x^*),$$

*where $\lambda^*, \nu^*$ is a global optimal solution of the dual.*

### Proof.

Since $\lambda \geq 0$, $g(x^*) \leq 0$ and $h(x^*) = 0$ it follows that

$$\lambda^T g(x^*) + \nu^T h(x^*) \leq 0.$$

On the other hand, for any $\lambda \geq 0, \nu$ and $x^* \in X$ :

$$d(\lambda, \nu) \leq f(x^*) + \lambda^T g(x^*) + \nu^T h(x^*),$$

and hence $d(\lambda, \nu) \leq \max_{\lambda \geq 0, \nu} d(\lambda, \nu) = d(\lambda^*, \nu^*) \leq f(x^*).$

## Duality in nonlinear optimization

If $d(\lambda^*, \nu^*) < f(x^*)$, then the difference $f(x^*) - d(\lambda^*, \nu^*)$ is called the duality gap.

The following result is an immediate consequence of last theorem.

**Theorem**
*A point $(x^*, \lambda^*, \nu^*) \in X \times \mathbb{R}^m_+ \times \mathbb{R}^p$ is a saddlepoint of the Lagrangian $L(x, \lambda, \nu)$ iff $x^*$ is a global minimum point of the primal problem (P), $\lambda^*, \nu^*$ is a global maximum of the dual (D), and the optimal values $f(x^*)$ of (P) and $d(\lambda^*, \nu^*)$ of (D) coincide.*

**Theorem**
*A point $(x^*, \lambda^*, \nu^*) \in X \times \mathbb{R}^m_+ \times \mathbb{R}^p$ is a saddlepoint of the Lagrangian $L(x, \lambda, \nu)$ iff the following conditions hold:*

$L(x^*, \lambda^*, \nu^*) = \min\{L(x, \lambda^*, \nu^*) : x \in X\}$.

$g(x^*) \leq 0$, $h(x^*) = 0$, $\lambda^{*T} g(x^*) = 0$.



## Duality in nonlinear optimization

**Example**
Consider the quadratic programming problem

$$\begin{array}{ll} \min & f(x) = c^T x + \frac{1}{2} x^T Q x \\ \text{s.t.} & Ax \leq b, \end{array}$$

where the $(n \times n)$–matrix $Q$ is symmetric positive definite, $c \in \mathbb{R}^n$, $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^n$.

The corresponding Lagrangian function is

$$\begin{aligned} L(x, \lambda) &= c^T x + \tfrac{1}{2} x^T Q x + \lambda^T (Ax - b) \\ &= (c + A^T \lambda)^T x + \tfrac{1}{2} x^T Q x - b^T \lambda \end{aligned}$$

The minimum of $L(x, \lambda)$ with respect to $x$ occurs at the point $x^*$ where $L'(x^*, \lambda) = 0$, i.e.,

$$x^* = -Q^{-1}(c + A^T \lambda).$$



## Duality in nonlinear optimization

Substituting in $L$ we obtain the dual function

$$d(\lambda) = -\frac{1}{2} \lambda^T A Q^{-1} A^T \lambda - \lambda^T (b + A Q^{-1} c) - \frac{1}{2} c^T Q^{-1} c.$$

Hence the dual problem is given by

$$\max_{\lambda \geq 0} d(\lambda) = -\frac{1}{2} \lambda^T M \lambda + d^T \lambda,$$

where $M = A Q^{-1} A^T$ and $d = -(b + A Q^{-1} c)$.

If $\lambda^*$ is the solution of the dual problem, then

$$x^* = -Q^{-1}(c + A^T \lambda^*)$$

is the solution of the original primal problem and

$$f(x^*) = d(\lambda^*).$$



## Primal-dual interior point methods

Consider the problem in the form:

$$\begin{array}{ll} \min & f_0(x), \\ \text{s.t.} & f_i(x) \leq 0, \; i = 1, \ldots, m, \\ & Ax = b, \end{array}$$

where $f_i : \mathbb{R}^n \to \mathbb{R}$, $i = 1, \ldots, m$, are convex twice continuously differentiable functions and $A \in \mathbb{R}^{p \times n}$ with rank $p < n$. We assume that the problem is strictly feasible.

Solving this problem is equivalent to solving the KKT system

$$\begin{aligned} f_0'(x) + \sum_{i=1}^m \lambda_i f_i'(x) + A^T \nu &= 0 \\ \lambda_i f_i(x) &= 0, \; i = 1, \ldots, m, \\ Ax &= b, \\ f_i(x) &\leq 0, \; i = 1, \ldots, m, \\ \lambda &\geq 0. \end{aligned}$$

## Primal-dual interior point methods

Denote by $\phi(x) = -\sum_{i=1}^{m} \ln(-f_i(x))$ the standard logarithmic barrier and consider the problem

$$\begin{array}{ll} \min & tf_0(x) + \phi(x), \\ \text{s.t.} & Ax = b. \end{array}$$

Then the central path is given by $x^*(t)$ such that

$$Ax^*(t) = b,\ f_i(x^*(t)) < 0, i = 1, \dots, m,$$

and for some $\hat{\nu} \in \mathbb{R}^p$

$$\begin{aligned} 0 &= tf_0'(x^*(t)) + \phi'(x^*(t)) + A^T\hat{\nu} \\ &= tf_0'(x^*(t)) + \sum_{i=1}^{m} \tfrac{1}{-f_i(x^*(t))} f_i'(x^*(t)) + A^T\hat{\nu}. \end{aligned}$$



## Primal-dual interior point methods

Denoting by

$$\lambda_i^*(t) = -\frac{1}{tf_i(x^*(t))}, i = 1, \dots, m;\ \nu^*(t) = \hat{\nu}/t,$$

we obtain

$$f_0'(x^*(t)) + \sum_{i=1}^{m} \lambda_i^*(t) f_i'(x^*(t)) + A^T\nu^*(t) = 0.$$

Thus, $x^*(t)$ minimizes the Lagrangian

$$L(x, \lambda, \nu) = f_0(x) + \sum_{i=1}^{m} \lambda_i f_i(x) + \nu^T(Ax - b)$$

for $\lambda = \lambda^*(t)$ and $\nu = \nu^*(t)$. Then $\lambda^*(t), \nu^*(t)$ is dual feasible and

$$\begin{aligned} g(\lambda^*(t), \nu^*(t)) &= f_0(x^*(t)) + \sum_{i=1}^{m} \lambda_i^*(t) f_i(x^*(t)) + \nu^*(t)^T(Ax^*(t) - b) \\ &= f_0(x^*(t)) - m/t. \end{aligned}$$

Hence, $f_0(x^*(t)) - f_0^* \le m/t$.



## Primal-dual interior point methods

Thus, the central path conditions can be interpreted as "deformed KKT conditions"

$$\begin{aligned} f_0'(x^*(t)) + \sum_{i=1}^{m} \lambda_i^*(t) f_i'(x^*(t)) + A^T\nu^*(t) &= 0 \\ -\lambda_i^*(t) f_i(x^*(t)) &= 1/t,\ i = 1, \dots, m, \\ Ax^*(t) &= b, \\ f_i(x^*(t)) &\le 0,\ i = 1, \dots, m, \\ \lambda^*(t) &\ge 0, \end{aligned}$$

where the complementary slackness condition in KKT is replaced with

$$\lambda_i^*(t) f_i(x^*(t)) = 1/t.$$

The search directions of a primal-dual interior point method are obtained from Newton method applied to the modified KKT equations.



## Primal-dual interior point methods

We are dealing with both primal and dual variables. We express the modified KKT conditions as $r_t(x, \lambda, \nu) = 0$, where

$$r_t(x, \lambda, \nu) = \begin{bmatrix} f_0'(x) + Df(x)^T\lambda + A^T\nu \\ -\operatorname{diag}(\lambda) f(x) - (1/t)e \\ Ax - b \end{bmatrix},$$

$t = 0, e = (1, \dots, 1)^T \in \mathbb{R}^m$ and

$$f(x) = \begin{bmatrix} f_1(x) \\ \dots \\ f_m(x) \end{bmatrix},\ Df(x) = \begin{bmatrix} f_1'(x)^T \\ \dots \\ f_m'(x)^T \end{bmatrix}.$$

# Primal-dual interior point methods

Denote the current point and Newton step by

$$y = (x, \lambda, \nu), \ \Delta y = (\Delta x, \Delta\lambda, \Delta\nu).$$

Then the Newton step is obtained from the linear system

$$r_t(y + \Delta y) \approx r_t(y) + Dr_t(y)\Delta y = 0,$$

which is given by

$$\begin{bmatrix} f_0''(x) + \sum_{i=1}^{m} \lambda_i f_i''(x) & Df(x)^T & A^T \\ -\mathrm{diag}(\lambda) Df(x) & -\mathrm{diag}(f(x)) & 0 \\ A & 0 & 0 \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta\lambda \\ \Delta\nu \end{bmatrix} = - \begin{bmatrix} r_d \\ r_c \\ r_p \end{bmatrix},$$

where

$$\begin{bmatrix} r_d \\ r_c \\ r_p \end{bmatrix} = \begin{bmatrix} f_0'(x) + Df(x)^T\lambda + A^T\nu \\ -\mathrm{diag}(\lambda) f(x) - (1/t)e \\ Ax - b \end{bmatrix}.$$