ORF522 – Linear and Nonlinear Optimization

8. Linear optimization duality
Ed Forum

• Why do we need to solve dual problem instead of the primal problem? When we have a LP problem, in what scenario does solving dual problem more efficient than primal problem?

• How does the definition of $y$ imply nonnegative reduced costs?
Recap
Optimal objective values

Primal
minimize \( c^T x \)
subject to \( Ax \leq b \)

\( p^* \) is the primal optimal value

Dual
maximize \( -b^T y \)
subject to \( A^T y + c = 0 \)
\( y \geq 0 \)

\( d^* \) is the dual optimal value

Primal infeasible: \( p^* = +\infty \)
Primal unbounded: \( p^* = -\infty \)

Dual infeasible: \( d^* = -\infty \)
Dual unbounded: \( d^* = +\infty \)
Relationship between primal and dual

<table>
<thead>
<tr>
<th>$d^* = +\infty$</th>
<th>$p^* = +\infty$</th>
<th>$p^*$ finite</th>
<th>$p^* = -\infty$</th>
</tr>
</thead>
<tbody>
<tr>
<td>primal inf. dual unb.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$d^*$ finite</td>
<td></td>
<td>optimal values equal</td>
<td></td>
</tr>
<tr>
<td>$d^* = -\infty$</td>
<td>exception</td>
<td></td>
<td>primal unbounded dual inf</td>
</tr>
</tbody>
</table>

- Upper-right excluded by **weak duality**
- $(1, 1)$ and $(3, 3)$ proven by **weak duality**
- $(3, 1)$ and $(2, 2)$ proven by **strong duality**
Today's agenda
Readings: [Chapter 4, LO][Chapter 11, LP]

• Two-person zero-sum games
• Farkas lemma
• Complementary slackness
• Dual simplex method
Two-person zero-sum games
Rock paper scissors

Rules
At count to three declare one of: Rock, Paper, or Scissors

Winners
Identical selection is a draw, otherwise:
• Rock beats (“dulls”) scissors
• Scissors beats (“cuts”) paper
• Paper beats (“covers”) rock

Extremely popular: world RPS society, USA RPS league, etc.
Two-person zero-sum game

• Player 1 (P1) chooses a number $i \in \{1, \ldots, m\}$ (one of $m$ actions)
• Player 2 (P2) chooses a number $j \in \{1, \ldots, n\}$ (one of $n$ actions)

Two players make their choice independently

Rule

Player 1 pays $A_{ij}$ to player 2

$A \in \mathbb{R}^{m \times n}$ is the payoff matrix

Rock, Paper, Scissors

$$
\begin{bmatrix}
R & P & S \\
R & 0 & 1 & -1 \\
A = & P & -1 & 0 & 1 \\
S & 1 & -1 & 0
\end{bmatrix}
$$
Mixed (randomized) strategies

Deterministic strategies can be systematically defeated

Randomized strategies
- P1 chooses randomly according to distribution $x$:
  \[ x_i = \text{probability that P1 selects action } i \]
- P2 chooses randomly according to distribution $y$:
  \[ y_i = \text{probability that P2 selects action } j \]

Expected payoff (from P1 P2), if they use mixed-strategies $x$ and $y$,

\[
\sum_{i=1}^{m} \sum_{j=1}^{n} x_i y_j A_{ij} = x^T A y
\]
Mixed strategies and probability simplex

**Probability simplex** in $\mathbb{R}^k$

$$P_k = \{ p \in \mathbb{R}^k \mid p \geq 0, \quad 1^T p = 1 \}$$

**Mixed strategy**

For a game player, a mixed strategy is a distribution over all possible deterministic strategies.

The **set of all mixed strategies** is the probability simplex $\rightarrow x \in P_m, \quad y \in P_n$
Optimal mixed strategies

P1: optimal strategy $x^*$ is the solution of

\[
\text{minimize } \max_{y \in P_n} x^T Ay \quad \text{subject to } x \in P_m
\]

\[
\text{minimize } \max_{j=1,\ldots,n} (A^T x)_j \quad \text{subject to } x \in P_m
\]

Inner problem over deterministic strategies (vertices)

P2: optimal strategy $y^*$ is the solution of

\[
\text{maximize } \min_{x \in P_m} x^T Ay \quad \text{subject to } y \in P_n
\]

\[
\text{maximize } \min_{i=1,\ldots,m} (Ay)_i \quad \text{subject to } y \in P_n
\]

Optimal strategies $x^*$ and $y^*$ can be computed using linear optimization.
Minmax theorem

Theorem

\[
\max_{y \in P_n} \min_{x \in P_m} x^T A y = \min_{x \in P_m} \max_{y \in P_n} x^T A y
\]

Proof

The optimal \( x^* \) is the solution of

- minimize \( t \)
- subject to \( A^T x \leq t \mathbf{1} \)
- \( \mathbf{1}^T x = 1 \)
- \( x \geq 0 \)

The optimal \( y^* \) is the solution of

- maximize \( w \)
- subject to \( Ay \geq w \mathbf{1} \)
- \( \mathbf{1}^T y = 1 \)
- \( y \geq 0 \)

The two LPs are duals and by strong duality the equality follows.
Nash equilibrium

Theorem

\[
\max_{y \in P_n} \min_{x \in P_m} x^T Ay = \min_{x \in P_m} \max_{y \in P_n} x^T Ay
\]

Consequence

The pair of mixed strategies \((x^*, y^*)\) attains the Nash equilibrium of the two-person matrix game, i.e.,

\[
x^T Ay^* \geq x^{*T} Ay^* \geq x^T Ay, \quad \forall x \in P_m, \forall y \in P_n
\]
Example

\[
A = \begin{bmatrix}
4 & 2 & 0 & -3 \\
-2 & -4 & -3 & 3 \\
-2 & -3 & 4 & 1
\end{bmatrix}
\]

\[
\min_i \max_j A_{ij} = 3 > -2 = \max_j \min_i A_{ij}
\]

Optimal mixed strategies

\[
x^* = (0.37, 0.33, 0.3), \quad y^* = (0.4, 0, 0.13, 0.47)
\]

Expected payoff

\[
x^T Ay^* = 0.2
\]
Farkas lemma
Feasibility of polyhedra

\[ P = \{ x \mid Ax = b, \quad x \geq 0 \} \]

How to show that \( P \) is \textbf{feasible}?
Easy: we just need to provide an \( x \in P \), i.e., a \textbf{certificate}

How to show that \( P \) is \textbf{infeasible}?
Farkas lemma

Theorem
Given $A$ and $b$, exactly one of the following statements is true:

1. There exists an $x$ with $Ax = b$, $x \geq 0$
2. There exists a $y$ with $A^T y \geq 0$, $b^T y < 0$
Farkas lemma

Geometric interpretation

1. First alternative
   There exists an \( x \) with \( Ax = b, \ x \geq 0 \)
   
   \[
   b = \sum_{i=1}^{n} x_i A_i, \quad x_i \geq 0, \ i = 1, \ldots, n
   \]
   
   \( b \) is in the cone generated by the columns of \( A \)

2. Second alternative
   There exists a \( y \) with \( A^T y \geq 0, \ b^T y < 0 \)
   
   \[
   y^T A_i \geq 0, \quad i = 1, \ldots, m, \quad y^T b < 0
   \]
   
   The hyperplane \( y^T z = 0 \) separates \( b \) from \( A_1, \ldots, A_n \)
Farkas lemma

There exists \( x \) with \( Ax = b, \ x \geq 0 \) \hspace{1cm} \text{OR} \hspace{1cm} \text{There exists } y \text{ with } A^T y \geq 0, \ b^T y < 0

Proof

1 and 2 cannot be both true (easy)

\[
x \geq 0, \ Ax = b \text{ and } y^T A \geq 0 \quad \rightarrow \quad y^T b = y^T Ax \geq 0
\]
Farkas lemma

There exists $x$ with $Ax = b$, $x \geq 0$ \hspace{1cm} \text{OR} \hspace{1cm} \text{There exists } y \text{ with } A^T y \geq 0, \ b^T y < 0$

Proof

1 and 2 cannot be both false (duality)

<table>
<thead>
<tr>
<th>Primal</th>
<th>Dual</th>
</tr>
</thead>
<tbody>
<tr>
<td>minimize $0$</td>
<td>maximize $-b^T y$</td>
</tr>
<tr>
<td>subject to $Ax = b$</td>
<td>subject to $A^T y \geq 0$</td>
</tr>
<tr>
<td>$x \geq 0$</td>
<td></td>
</tr>
</tbody>
</table>

$y = 0$ always feasible

Strong duality holds

$d^* \neq -\infty, \ p^* = d^*$
Farkas lemma

There exists \( x \) with \( Ax = b, \ x \geq 0 \) \hspace{1cm} \text{OR} \hspace{1cm} \text{There exists} \ y \text{ with } A^T y \geq 0, \ b^T y < 0

Proof

1 and 2 cannot be both false (duality)

<table>
<thead>
<tr>
<th>Primal</th>
<th>Dual</th>
</tr>
</thead>
<tbody>
<tr>
<td>minimize 0</td>
<td>maximize (-b^T y)</td>
</tr>
<tr>
<td>subject to ( Ax = b )</td>
<td>subject to ( A^T y \geq 0 )</td>
</tr>
<tr>
<td>( x \geq 0 )</td>
<td>( x \geq 0 )</td>
</tr>
</tbody>
</table>

Alternative 1: primal feasible \( p^* = d^* = 0 \)

\( b^T y \geq 0 \) for all \( y \) such that \( A^T y \geq 0 \)
Farkas lemma

There exists $x$ with $Ax = b$, $x \geq 0$ \quad \text{OR} \quad \text{There exists } y \text{ with } A^T y \geq 0, b^T y < 0

Proof

1 and 2 cannot be both false (duality)

<table>
<thead>
<tr>
<th>Primal</th>
<th>Dual</th>
</tr>
</thead>
<tbody>
<tr>
<td>minimize 0</td>
<td>maximize $-b^T y$</td>
</tr>
<tr>
<td>subject to $Ax = b$</td>
<td>subject to $A^T y \geq 0$</td>
</tr>
<tr>
<td>$x \geq 0$</td>
<td>$x \geq 0$</td>
</tr>
</tbody>
</table>

Alternative 2: primal infeasible $p^* = d^* = +\infty$

There exists $y$ such that $A^T y \geq 0$ and $b^T y < 0$
Farkas lemma

Many variations

There exists $x$ with $Ax = b$, $x \geq 0$

**OR**

There exists $y$ with $A^T y \geq 0$, $b^T y < 0$

There exists $x$ with $Ax \leq b$, $x \geq 0$

**OR**

There exists $y$ with $A^T y \geq 0$, $b^T y < 0$, $y \geq 0$

There exists $x$ with $Ax \leq b$

**OR**

There exists $y$ with $A^T y = 0$, $b^T y < 0$, $y \geq 0$
Complementary slackness
Optimality conditions

<table>
<thead>
<tr>
<th>Primal</th>
<th>Dual</th>
</tr>
</thead>
<tbody>
<tr>
<td>minimize $c^T x$</td>
<td>maximize $-b^T y$</td>
</tr>
</tbody>
</table>
| subject to $Ax \leq b$  | subject to $A^T y + c = 0$  
|                         | $y \geq 0$          |

$x$ and $y$ are **primal** and **dual** optimal if and only if

- $x$ is **primal feasible**: $Ax \leq b$
- $y$ is **dual feasible**: $A^T y + c = 0$ and $y \geq 0$
- The **duality gap** is zero: $c^T x + b^T y = 0$

Can we relate $x$ and $y$ (not only the objective)?
Complementary slackness

<table>
<thead>
<tr>
<th>Primal</th>
<th>Dual</th>
</tr>
</thead>
<tbody>
<tr>
<td>minimize $c^T x$</td>
<td>maximize $-b^T y$</td>
</tr>
<tr>
<td>subject to $Ax \leq b$</td>
<td>subject to $A^T y + c = 0$</td>
</tr>
</tbody>
</table>

**Theorem**

Primal, dual feasible $x, y$ are optimal if and only if

$$y_i(b_i - a_i^T x) = 0, \quad i = 1, \ldots, m$$

i.e., at optimum, $b - Ax$ and $y$ have a **complementary sparsity** pattern:

$$y_i > 0 \quad \Rightarrow \quad a_i^T x = b_i$$

$$a_i^T x < b_i \quad \Rightarrow \quad y_i = 0$$
Complementary slackness

**Proof**

The duality gap at primal feasible $x$ and dual feasible $y$ can be written as

$$c^T x + b^T y = (-A^T y)^T x + b^T y = (b - Ax)^T y = \sum_{i=1}^{m} y_i (b_i - a_i^T x) = 0$$

Since all the elements of the sum are nonnegative, they must all be 0

For feasible $x$ and $y$ complementary slackness = zero duality gap
Geometric interpretation

Example in $\mathbb{R}^2$

Two active constraints at optimum: $a_1^T x^* = b_1$, $a_2^T x^* = b_2$

Optimal dual solution $y$ satisfies:

$$A^T y + c = 0, \quad y \geq 0, \quad y_i = 0 \text{ for } i \neq \{1, 2\}$$

In other words, $-c = a_1 y_1 + a_2 y_2$ with $y_1, y_2 \geq 0$

**Geometric interpretation:** $-c$ lies in the cone generated by $a_1$ and $a_2$
Example

minimize \(-4x_1 - 5x_2\)

subject to

\[
\begin{bmatrix}
-1 & 0 \\
2 & 1 \\
0 & -1 \\
1 & 2
\end{bmatrix}
\begin{bmatrix}
x_1 \\
x_2
\end{bmatrix}
\leq
\begin{bmatrix}
0 \\
3 \\
0 \\
3
\end{bmatrix}
\]

Let’s show that feasible \(x = (1, 1)\) is optimal

Second and fourth constraints are active at \(x \rightarrow y = (0, y_2, 0, y_4)\)

\[
A^T y = -c \quad \Rightarrow \quad \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} y_2 \\ y_4 \end{bmatrix} = \begin{bmatrix} 4 \\ 5 \end{bmatrix}
\]

and

\(y_2 \geq 0, \quad y_4 \geq 0\)

\(y = (0, 1, 0, 2)\) satisfies these conditions and proves that \(x\) is optimal

Complementary slackness is useful to recover \(y^*\) from \(x^*\)
The dual simplex
Primal and dual basic feasible solutions

**Primal problem**
- **minimize** \( c^T x \)
- **subject to** \( Ax = b \)
- \( x \geq 0 \)

**Dual problem**
- **maximize** \(-b^T y\)
- **subject to** \( A^T y + c \geq 0\)

Given a **basis** matrix \( A_B \)

**Primal feasible:** \( Ax = b, \ x \geq 0 \quad \Rightarrow \quad x_B = A_B^{-1}b \geq 0\)

**Dual feasible:** \( A^T y + c \geq 0\).

If \( y = -A_B^{-T}c_B \quad \Rightarrow \quad c - A^T A_B^{-T}c_B \geq 0\)

**Zero duality gap:** \( c^T x + b^T y = c_B^T x_B - b^T A_B^{-T}c_B = c_B^T x_B - c_B^T A_B^{-1}b = 0\) (by construction)
The primal (dual) simplex method

**Primal problem**
- minimize $c^T x$
- subject to $Ax = b$ \( x \geq 0 \)

**Dual problem**
- maximize $-b^T y$
- subject to $A^T y + c \geq 0$

**Primal simplex**
- Primal feasibility
- Zero duality gap

**Dual feasibility**

**Dual simplex**
- Dual feasibility
- Zero duality gap

**Primal feasibility**
Feasible dual directions

Conditions

\[ P = \{ y \mid A^T y + c \geq 0 \} \]

Given a basis matrix \( A_B = \begin{bmatrix} A_{B(1)} & \cdots & A_{B(m)} \end{bmatrix} \)
we have dual feasible solution \( y \):

\[ \bar{c} = A^T y + c \geq 0 \]

Feasible direction \( d \)

\[ y + \theta d \]

Reduced cost change

\[ c + A^T (y + \theta d) \geq 0 \quad \Rightarrow \quad \bar{c} + \theta z \geq 0 \]

\[ A^T d = z \quad \text{(subspace restriction)} \]
Feasible directions

**Computation**

\[ \bar{c} + \theta z \geq 0 \]
\[ A^T d = z \]

**Subspace restriction**

\[ A_B^T d = z_B \]
\[ A_N^T d = z_N \]

**Basic indices**

\[ z_B = e_i \quad \rightarrow \quad B(\ell) = i \text{ exits the basis} \]

Get \( d \) by solving \( A_B^T d = z_B \)

**Nonbasic indices**

\[ z_N = A_N^T d = A_N^T A_B^{-T} e_i \]

**Non-negativity of reduced costs (non-degenerate assumption)**

- Basic variables: \( \bar{c}_B = 0 \). Nonnegative direction \( z_B \geq 0 \).
- Nonbasic variables: \( \bar{c}_N > 0 \). Therefore \( \exists \theta > 0 \) such that \( \bar{c}_N + \theta z_N \geq 0 \)
Stepsizes

How far can we go?

\[ \theta^* = \max\{ \theta \mid \theta \geq 0 \text{ and } \bar{c} + \theta z \geq 0 \} \]

Unbounded

If \( z \geq 0 \), then \( \theta^* = \infty \). The dual problem is unbounded (primal infeasible).

Bounded

If \( z_j < 0 \) for some \( j \), then

\[ \theta^* = \min_{\{j \mid z_j < 0\}} \left( -\frac{\bar{c}_j}{z_j} \right) = \min_{\{j \in N \mid z_j < 0\}} \left( -\frac{\bar{c}_j}{z_j} \right) \]

(Since \( z_j \geq 0 \), \( j \in B \))
Moving to a new basis

Next reduced cost

\[
\bar{c} + \theta^* z
\]

Let \( j \notin \{B(1), \ldots, B(m)\} \) be the index such that \( \theta^* = -\frac{\bar{c}_j}{z_j} \). Then,

\[
\bar{c}_j + \theta^* z_j = 0
\]

New basis

\[
A_{\bar{B}} = \begin{bmatrix} A_{B(1)} & \cdots & A_{B(\ell-1)} & A_j & A_{B(\ell+1)} & \cdots & A_{B(m)} \end{bmatrix}
\]

New solution

\[
A_{\bar{B}} x_{\bar{B}} = b
\]
An iteration of the dual simplex method

**Initialization**

- a basic dual feasible solution $y$, i.e. $A^T y + c \geq 0$
- a basis matrix $A_B = \begin{bmatrix} A_B(1) & \cdots & A_B(m) \end{bmatrix}$

**Remark**

Reduced costs nonnegative

**Iteration steps**

1. Get $x$
   - Solve $A_B x_B = b \left( O\left( m^2 \right) \right)$
   - Set $x_i = 0$ if $i \notin B$

2. If $x \geq 0$, $x$ feasible. break

3. Choose $i$ such that $x_i < 0$

4. Compute each direction $z$ with
   - $z_i = 1$, $A_B^T d = e_i$ and $z_N = A_N^T d \left( O\left( m^2 \right) \right)$

5. If $z_N \geq 0$, the dual problem is unbounded and the optimal value is $+\infty$. break

6. Compute step length $\theta^* = \min_{\{j \in N | z_j < 0\}} \left( \frac{-c_j}{z_j} \right)$

7. Compute new point $y + \theta^* d$

8. Get new basis $A_B = A_B + (A_j - A_i)e_\ell^T$
   - perform rank-1 factor update ($j$ enters, $i$ exists) $O\left( m^2 \right)$
**Example**

From lecture 6

minimize \( c^T x \)
subject to \( Ax = b \)
\( x \geq 0 \)

Dual problem

maximize \(-b^T y\)
subject to \( A^T y + c \geq 0 \)

Initialize

\( y = (10, 0, 0) \)
\( B = \{1, 5, 6\} \)
\n\( c + A^T y = (0, 8, 8, 10, 0, 0) \geq 0 \)
Example

Iteration 1

Primal solution \( x = (20, 0, 0, 0, -20, -20) \)

Solve \( Ax_B = b \) \( \Rightarrow \) \( x_B = (20, -20, -20) \)

Direction \( z = (0, -3, -2, -2, 1, 0), \quad i = 5 \)

Solve \( A^T_B d = e_i \) \( \Rightarrow \) \( d = (-2, 1, 0) \)

Get \( z_N = A^T_N d = (-3, -2, -2) \)

Step \( \theta^* = 2.66, \quad j = 2 \)

\( \theta^* = \min \left( -\frac{\bar{c}_j}{z_j} \right) = \{2.66, 4, 5\} \)

New \( y \leftarrow y + \theta^* d = (4.66, 2.66, 0) \)
Example

Iteration 2

Primal solution \( x = (6.66, 6.66, 0, 0, 0, -6.66) \)
Solve \( Ax_B = b \) \( \Rightarrow \) \( x_B = (6.66, 6.66, -6.66) \)

Direction \( z = (0, 0, -1.66, -0.66, -0.66, 1), \quad i = 6 \)
Solve \( A_B^T d = e_i \) \( \Rightarrow \) \( d = (-0.66, -0.66, 1) \)
Get \( z_N = A_N^T d = (-1.66, -0.66, -0.66) \)

Step \( \theta^* = 1.6, \quad j = 3 \)
\( \theta^* = \min_{\{j \mid z_j < 0\}} (-\bar{c}_j/z_j) = \{1.6, 7, 4\} \)
New \( y \leftarrow y + \theta^* d = (3.6, 1.6, 1.6) \)
Example
Iteration 3

\[ y = (3.6, 1.6, 1.6) \]
\[ -b^T y = -136 \]
\[ c + A^T y = (0, 0, 0, 3.6, 1.6, 1.6) \]
\[ B = \{1, 2, 3\} \]
\[ A_B = \begin{bmatrix}
1 & 2 & 2 \\
2 & 1 & 2 \\
2 & 2 & 1
\end{bmatrix} \]

Primal solution \[ x = (4, 4, 4, 0, 0, 0) \]
Solve \[ A x_B = b \quad \Rightarrow \quad x_B = (4, 4, 4) \]

\[ x \geq 0 \quad \rightarrow \quad x^* = (4, 4, 4, 0, 0, 0) \]

Optimal solution

Same as primal simplex!

\[ c = (-10, -12, -12, 0, 0, 0) \]
\[ A = \begin{bmatrix}
1 & 2 & 2 & 1 & 0 & 0 \\
2 & 1 & 2 & 0 & 1 & 0 \\
2 & 2 & 1 & 0 & 0 & 1
\end{bmatrix} \]
\[ b = (20, 20, 20) \]
Equivalence and symmetry

The dual simplex is equivalent to the primal simplex applied to the dual problem.

Dual problem
maximize \(-b^T y\)
subject to \(A^T y + c \geq 0\)

Standard form
minimize
\[
\begin{bmatrix}
  b & -b & 0
\end{bmatrix} w
\]
subject to
\[
\begin{bmatrix}
  A^T & -A^T & -I
\end{bmatrix} w = -c
\]
\(w \geq 0\)
\(w = (y^+, y^-, s)\)
Dual simplex efficiency

Sequence of problems with varying feasible region

previous $y$ still dual feasible $\longrightarrow$ **warm-start**

Applied in many different contexts, for example:

1. **sequential decision-making**
2. **mixed-integer optimization** to solve subproblems

(more later in the course…)
Linear optimization duality

Today, we learned to:

• **Interpret** linear optimization duality using game theory
• **Prove** Farkas lemma using duality
• **Geometrically link** primal and dual solutions with complementary slackness
• **Implement** the dual simplex method
Next lecture

• Sensitivity analysis