Optimization Techniques for Machine Learning

AMLZC326 · #04 Linear Optimization II

Anshid Aboobacker

Table of Contents

MOTIVATION FOR ARTIFICIAL VARIABLES

- **Standard Simplex Requirements:**
  - Linear Programs (LPs) with $(\leq)$ constraints and non-negative RHS offer a convenient **all-slack starting basic feasible solution**.
- **The Problem:**
  - Models involving $(=)$ and/or $(\geq)$ constraints do not possess an automatic starting identity matrix.
  - These are termed ”ill-behaved” LPs.
- **The Solution:**
  - Use **Artificial Variables** to play the role of slacks for the first iteration.
  - These variables are disposed of at a later iteration.

LEARNING OBJECTIVES

By the end of this lecture you should be able to:

- Handle LPPs lacking an obvious initial BFS using the Big-M method and the Two-Phase method
- Perform sensitivity analysis: determine how the optimal solution changes with problem data (RHS and objective coefficients)
- Construct the dual LP for a given primal and interpret shadow prices as marginal values of constraints
- State the strong duality result: at optimality, primal objective equals dual objective

Methods for Artificial Variables

Two closely related methods are used to handle artificial variables:

1. **The M-Method:** Uses a large penalty in the objective function.
2. **The Two-Phase Method:** Separates the search for feasibility (Phase I) and optimality (Phase II).

The M-Method: Concept

- Start with the LP in equation form.
- If equation i lacks a slack variable, add an artificial variable R_i.
- **The Modeling "Trick":**
  - Artificial variables are not part of the original problem.
  - We must force them to zero by the optimum iteration.
- **The Penalty:** Assign a penalty M in the objective function.

$$\text{Coef. of } R_i = \begin{cases} -M & \text{in maximization problems} \\ +M & \text{in minimization problems} \end{cases}$$

- M represents a sufficiently large positive value ($M \to \infty$).

Example

Minimize $z = 4x_1 + x_2$

Subject to:

$$\begin{aligned} 3x_1 + x_2 &= 3 \\ 4x_1 + 3x_2 &\geq 6 \\ x_1 + 2x_2 &\leq 4 \\ x_1, x_2 &\geq 0 \end{aligned}$$

Convert to Equation Form

- Constraint 2: Subtract surplus variable x_3.
- Constraint 3: Add slack variable x_4.

Add Artificial Variables

- First and second equations lack basic variables.
- Add artificial variables R_1 and R_2.

Penalized Objective Function: Minimize $z = 4x_1 + x_2 + MR_1 + MR_2$

EXAMPLE

Resulting System:

$$
\begin{aligned}
3x_1 + x_2 + R_1 &= 3 \\
4x_1 + 3x_2 - x_3 + R_2 &= 6 \\
x_1 + 2x_2 + x_4 &= 4
\end{aligned}
$$

Starting Basic Solution: $(R_1, R_2, x_4) = (3, 6, 4)$

Computational Note: Value of M

- Substitute sufficiently large numeric value for M. Must be large relative to original coefficients to force artificials to zero. Should not be excessively large to avoid roundoff error.
- Solve it as a Maximisation problem by taking $w = -z$
- Objective coefficients are -4 and -1. We set $M = -100$.

Initial Tableau & Consistency

Basic	w	x_1	x_2	x_3	x_4	R_1	R_2	Sol
w	1	4	1	0	0	100	100	0
R_1	0	3	1	0	0	1	0	3
R_2	0	4	3	-1	0	0	1	6
x_4	0	1	2	0	1	0	0	4

Initial Inconsistency:

- The basic variables R_1 and R_2 have non-zero coefficients (100) in the objective row.
- Current w-row shows $w = 0$, but actual cost is $-100(3) + -100(6) = -900$.

Correction Step: Substitute out R_1 and R_2 in the z-row:

$$\text{New } w\text{-row} = \text{Old } w\text{-row} + (-100 \times R_1\text{-row} + -100 \times R_2\text{-row})$$

Solving the Problem

Iteration 1

Basic	w	x_1	x_2	x_3	x_4	R_1	R_2	Sol
w	1	-696	-401	100	0	0	0	-900
R_1	0	3	1	0	0	1	0	3
R_2	0	4	3	-1	0	0	1	6
x_4	0	1	2	0	1	0	0	4

Canonical BFS: $(R_1, R_2, x_4) = (3, 6, 4)$
$w = -900$ **Iteration Logic:**

- **Enter:** x_1 (Most negative reduced cost: -696).
- **Leave:** R_1 (Minimum ratio test: $3/3 = 1$ is smallest).
- Objective increases in each iteration
- Artificial variables R_1 and R_2 leave the basis in the first two iterations.

ITERATIONS

Iteration 2

Basic	w	x_1	x_2	x_3	x_4	R_1	R_2	Sol
w	1	0	-167	100	0	232	0	-204
x_1	0	1	0.333333	0	0	0.333333	0	1
R_2	0	0	1.666667	-1	0	-1.333333	1	2
x_4	0	0	1.666667	0	1	-0.333333	0	3

BFS: $(x_1, R_2, x_4) = (1, 2, 3) \quad w = -204$

Iteration 3

Basic	w	x_1	x_2	x_3	x_4	R_1	R_2	Sol
w	1	0	0	-0.2	0	98.4	100.2	-3.6
x_1	0	1	0	0.2	0	0.6	-0.2	0.6
x_2	0	0	1	-0.6	0	-0.8	0.6	1.2
x_4	0	0	0	1	1	1	-1	1

BFS: $(x_1, x_2, x_4) = (0.6, 1.2, 1) \quad w = -3.6$

ITERATIONS

Iteration 4

Basic	w	x_1	x_2	x_3	x_4	R_1	R_2	Sol
w	1	0	0	0	0.2	98.6	100	-3.4
x_1	0	1	0	0	-0.2	0.4	0	0.4
x_2	0	0	1	0	0.6	-0.2	0	1.8
x_3	0	0	0	1	1	1	-1	1

BFS: $(x_1, x_2, x_3) = (0.4, 1.8, 1) \quad w = -3.4$
This table is optimal and $z^* = -w^* = 3.4$

The Two-Phase Method

- **Motivation:** Eliminates the constant M to prevent computer roundoff error .
- **Phase I (Feasibility):**
  - Add artificial variables exactly as in M-method.
  - Minimize the **sum of artificial variables** ($w = \sum R_i$).
    [Or equivalently, maximize $w = -\sum R_i$]
  - If min $w > 0$, the problem is infeasible.
  - If min $w = 0$, proceed to Phase II.
- **Phase II (Optimality):**
  - Use the feasible solution from Phase I as the starting basis for the original problem.
  - Use the original objective function

Example

We will solve the same example.
Phase I Objective: Maximize $w = -R_1 - R_2$

Basic	w	x_1	x_2	x_3	x_4	R_1	R_2	Sol
w	1	0	0	0	0	1	1	0
R_1	0	3	1	0	0	1	0	3
R_2	0	4	3	-1	0	0	1	6
x_4	0	1	2	0	1	0	0	4

Initial Consistency Step: Similar to M-Method, R_1 and R_2 must be substituted out of the r-row:

$$\text{New } r\text{-row} = \text{Old } r\text{-row} + (1 \times R_1\text{-row} + 1 \times R_2\text{-row})$$

EXAMPLE

Iteration 1 – Corrected Table

Basic	w	x_1	x_2	x_3	x_4	R_1	R_2	Sol
w	1	-7	-4	1	0	0	0	-9
R_1	0	3	1	0	0	1	0	3
R_2	0	4	3	-1	0	0	1	6
x_4	0	1	2	0	1	0	0	4

Iteration 2

Basic	w	x_1	x_2	x_3	x_4	R_1	R_2	Sol
w	1	0	-1.666667	1	0	2.333333	0	-2
x_1	0	1	0.333333	0	0	0.333333	0	1
R_2	0	0	1.666667	-1	0	-1.333333	1	2
x_4	0	0	1.666667	0	1	-0.333333	0	3

EXAMPLE

Iteration 3

Basic	w	x_1	x_2	x_3	x_4	R_1	R_2	Sol
w	1	0	0	0	0	1	1	0
x_1	0	1	0	0.2	0	0.6	-0.2	0.6
x_2	0	0	1	-0.6	0	-0.8	0.6	1.2
x_4	0	0	0	1	1	1	-1	1

End of Phase I:

- Optimum achieved with $w = 0$.
- Basic Feasible Solution: $x_1 = 3/5, x_2 = 6/5, x_4 = 1$.

Transition to Phase II:

- Artificial variables have completed their mission.
- Eliminate R columns from the tableau.
- Restore original objective function: Minimize $z = 4x_1 + x_2$.

Example

Phase 2

Basic	w	x_1	x_2	x_3	x_4	Sol
w	1	4	1	0	0	0
x_1	0	1	0	0.2	0	0.6
x_2	0	0	1	-0.6	0	1.2
x_4	0	0	0	1	1	1

Adjusting the Tableau:

- The Phase I tableau is used as the start for Phase II.
- **Inconsistency:** Basic variables x_1 and x_2 have non-zero coefficients in the restored z-row.
- **Fix:** Substitute out basic variables using row operations:

$$\text{New } z\text{-row} = \text{Old } z\text{-row} + (4 \times x_1\text{-row} + 1 \times x_2\text{-row})$$

Next Step: Solve using standard simplex.

EXAMPLE

Iteration 4

Basic	w	x_1	x_2	x_3	x_4	Sol
w	1	0	0	-0.2	0	-3.6
x_1	0	1	0	0.2	0	0.6
x_2	0	0	1	-0.6	0	1.2
x_4	0	0	0	1	1	1

Iteration 5

Basic	w	x_1	x_2	x_3	x_4	Sol
w	1	0	0	0	0.2	-3.4
x_1	0	1	0	0	-0.2	0.4
x_2	0	0	1	0	0.6	1.8
x_3	0	0	0	1	1	1

This table is optimal and $z^* = -w^* = 3.4$

Table of Contents

Sensitivity Analysis

- Sensitivity analysis studies how the optimal solution changes when the model parameters (input data) change.
- **Purpose:** Parameters are often estimates; we need to know how ”sensitive” the optimum is to variations.
- **Two Main Types of Analysis:**
    1. **Feasibility:** Sensitivity to changes in resource availability (Right-Hand Side of constraints).
    2. **Optimality:** Sensitivity to changes in unit profit or cost (Objective Function coefficients).

Changes in RHS of constraints

Example: JOBCO Problem

Two products on two machines:

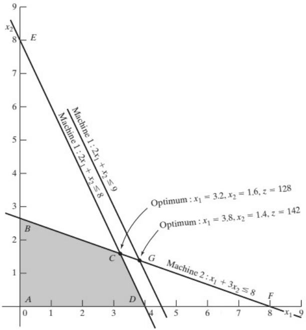
- Product 1: $30 revenue.
- Product 2: $20 revenue.

Visualizing Changes in Capacity

- Changing the Right-Hand Side (RHS) shifts the constraint line parallel to itself.
- **Example:** Increasing Machine 1 capacity from 8 to 9 hours.
- **Result:** The optimum point moves from *C* to *G* along the Machine 2 constraint.

Model:

$$
\begin{aligned}
\text{Maximize } z &= 30x_1 + 20x_2 \\
\text{s.t.} \quad 2x_1 + x_2 &\leq 8 \quad \text{(Machine 1)} \\
x_1 + 3x_2 &\leq 8 \quad \text{(Machine 2)} \\
x_1, x_2 &\geq 0
\end{aligned}
$$



Shadow Price/ Dual Price

The **Dual Price** (or Shadow Price) is the unit worth of a resource. It represents the rate of change of the objective function per unit change of a resource.

Old: $z = 128$ New: $z = 142$ Dual Price (M1) $= \frac{142-128}{9-8} = 14\$/\text{hr}$

Interpretation: 1 extra hour of M1 increases revenue by $14.

Feasibility Range: Dual Price valid only while optimal basis unchanged.
Min capacity (at B): $2(0) + 1(2.67) = 2.67$
Max capacity (at F): $2(8) + 1(0) = 16$

$$2.67 \leq \text{M1 Capacity} \leq 16$$

Within this range: Dual Price = $14/hr.

Similarly M2 Dual Price = $2/hr, with $4 \leq$ M2 Capacity ≤ 24.

Note: The dual prices equal the coefficients of the slack variables in the optimal z-row.

Economic Decision Making (Examples)

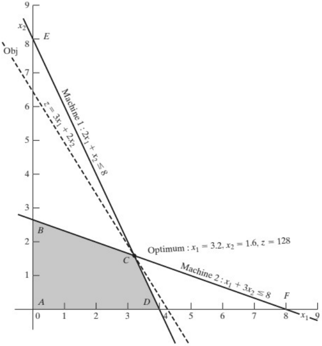
Using Dual Prices for Decision Making:

- **Q1:** If JOBCO can increase the capacity of both machines, which machine should receive priority?
  - Machine 1 value: \$14/hr. Machine 2 value: \$2/hr.
  - **Decision:** Prioritize Machine 1.
- **Q 2:** A suggestion is made to increase the capacities of machines 1 and 2 at the additional cost of \$10/hr for each machine. Is this advisable?
  - Proposal: Increase capacity at cost of \$10/hr.
  - Machine 1 Net: $14 - 10 = \$4$ (Profit).
  - Machine 2 Net: $2 - 10 = -\$8$ (Loss).
  - **Decision:** Only increase Machine 1.

Changes in Objective Coefficient

Graphical Insight:

- Changing (c_1, c_2) changes slope of objective line.
- Optimum C remains optimal if slope of z lies between slopes of the two binding constraints.



Model: Max $z = c_1 x_1 + c_2 x_2$

Binding constraints:

$x_1 + 3x_2 = 8, \qquad 2x_1 + x_2 = 8$

$$\frac{1}{3} \le \frac{c_1}{c_2} \le 2 \quad \left(0.333 \le \frac{c_1}{c_2} \le 2\right)$$

If outside this range: A new corner point becomes optimal.

Optimality Range (Examples)

Using Objective Coefficient Sensitivity:

Q1: Suppose unit revenues change to $c_1 = 35$, $c_2 = 25$. Will the current optimum remain the same?

- $\frac{c_1}{c_2} = \frac{35}{25} = 1.4$
- Since $0.333 \leq 1.4 \leq 2$, optimum at C remains unchanged.
- New objective value: $z = 35(3.2) + 25(1.6) = 152$
- **Conclusion:** Same corner point, different z.

Q2: If $c_2 = 20$, what range of c_1 keeps optimum unchanged?

- From $\frac{1}{3} \leq \frac{c_1}{c_2} \leq 2$ we get $\frac{1}{3} \leq \frac{c_1}{20} \leq 2$
- $6.67 \leq c_1 \leq 40$
- **Conclusion:** Within this range, (x_1, x_2) stay the same.

Table of Contents

Duality: Concept and Standard Form

Dual Problem: Defined systematically from the primal LPP.

- Optimal solution of one gives optimal solution of the other.
- Sometimes computationally easier to solve the dual.
- Provides economic interpretation and post-optimal analysis.

Standardized Approach to Duality

1. Write primal in equation form.
2. All RHS nonnegative.
3. All variables nonnegative.

This aligns with the simplex tableau and avoids memorizing multiple cases and rules.

Constructing the Dual Problem

Variables and Constraints:

- Assign a **dual variable** to each primal constraint.
- Assign a **dual constraint** to each primal variable.

Sign Relations:

- Primal equation $\rightarrow$ Dual variable unrestricted.
- Primal unrestricted variable $\rightarrow$ Dual equality constraint.

Coefficients:

- The **right-hand sides** of the primal constraints become the coefficients of the dual objective function.
- The primal objective coefficients become the **right-hand sides** of the dual constraints.
- The constraint coefficients are transposed (primal column becomes dual row).

Direction Rules:

Primal Problem	Dual Objective	Dual Constraints Type
Maximization	Minimization	$\geq$ (pointing up)
Minimization	Maximization	$\leq$ (pointing down)

Example

Primal: Max $z = 5x_1 + 12x_2 + 4x_3$

$$x_1 + 2x_2 + x_3 \leq 10$$
$$2x_1 - x_2 + 3x_3 = 8$$
$$x_1, x_2, x_3 \geq 0$$

Equation Form: Add slack x_4

$$x_1 + 2x_2 + x_3 + x_4 = 10$$
$$2x_1 - x_2 + 3x_3 + 0x_4 = 8$$

Dual: Min $w = 10y_1 + 8y_2$

$$y_1 + 2y_2 \geq 5 \quad (x_1)$$
$$2y_1 - y_2 \geq 12 \quad (x_2)$$
$$y_1 + 3y_2 \geq 4 \quad (x_3)$$
$$y_1 \geq 0 \quad (x_4)$$

y_2 unrestricted (primal equality constraint)

Relationship Between Objective Values

For any pair of feasible primal (max) and dual (min) solutions:

$$z = \sum c_j x_j \leq \sum b_i y_i = w$$

- **Weak Duality:** $z \leq w$ for any feasible solutions.
- **Strong Duality:** At the optimum, $z = w$.
- The optimum cannot occur with $z < w$.
- **Complementary Slackness (Max Form):** At optimum,
  $x_j \bar{c}_j = 0 \ \forall j \quad y_i(b_i - a_i x) = 0 \ \forall i$

Either a variable is zero, or its corresponding constraint is tight.

Example:
If $x_1 > 0$ at optimum $\Rightarrow \bar{c}_1 = 0$.
If constraint 1 has slack $(b_1 - a_1 x > 0) \Rightarrow y_1 = 0$.

Additional Simplex Methods (Big Picture)

1. **Primal Simplex**
   - Starts feasible.
   - Maintains feasibility.
   - Moves toward optimality.
2. **Dual Simplex**
   - Starts infeasible but satisfies optimality conditions.
   - Maintains optimality, seeks feasibility.
3. **Generalized Simplex**
   - If LP is infeasible and non-optimal:
   - Use Dual Simplex $\rightarrow$ restore feasibility.
   - Switch to Primal Simplex $\rightarrow$ reach optimality.

Key Takeaways

- Big-M method penalises artificial variables in the objective; Two-Phase method minimises their sum (Phase I) then optimises the original objective (Phase II)
- Sensitivity analysis: shadow price y_i^* is the rate of change of the optimal objective per unit relaxation of constraint i
- Every LP has a dual; at optimality: primal objective $=$ dual objective (strong duality)
- Complementary slackness: $x_i s_i = 0$ and $y_j t_j = 0$ — either a variable or its dual slack must be zero at the optimum

Thank you :)