

Mathematics Of Operations Research

Andrew Tulloch

Contents

Chapter 1. Generalities	5
Chapter 2. Constrained Optimization	6
1. Lagrangian Multipliers	6
2. Lagrange Dual	6
3. Supporting Hyperplanes	7
Chapter 3. Linear Programming	8
1. Convexity and Strong Duality	8
2. Linear Programs	9
3. Linear Program Duality	9
4. Complementary Slackness	9
Chapter 4. Simplex Method	11
1. Basic Solutions	11
2. Extreme Points and Optimal Solutions	11
3. The Simplex Tableau	11
4. The Simplex Method in Tableau Form	12
Chapter 5. Advanced Simplex Procedures	13
1. The Two-Phase Simplex Method	13
2. Gomory's Cutting Plane Method	13
Chapter 6. Complexity of Problems and Algorithms	14
1. Asymptotic Complexity	14
Chapter 7. The Complexity of Linear Programming	15
1. A Lower Bound for the Simplex Method	15
2. The Idea for a New Method	15
Chapter 8. Graphs and Flows	17
1. Introduction	17
2. Minimal Cost Flows	17

CONTENTS	4
Chapter 9. Transportation and Assignment Problems	18
Chapter 10. Non-Cooperative Games	19
Chapter 11. Strategic Equilibrium	20
Bibliography	21

CHAPTER 1

Generalities

www.statslab.cam.ac.uk/~ff271/teachin/mor ff271@cam.ac.uk

CHAPTER 2

Constrained Optimization

Minimize $f(x)$ subject to $h(x) = b, x \in X$.

Objective function $f : R^n \rightarrow R$ Vector $x \in R^n$ of decision variables, Functional constraint where $h : R^n \rightarrow R^m, b \in R^m$ Regional constraint where $X \subseteq R^n$.

DEFINITION 2.1. The feasible set is $X(b) = \{x \in X : h(x) = b\}$.

An inequality of the form $g(x) \leq b$ can be written as $g(x) + z = b$, where $z \in R^m$ called a slack variable with regional constraint $z \geq 0$.

1. Lagrangian Multipliers

DEFINITION 2.2. Define the Lagrangian of a problem as

$$L(x, \lambda) = f(x) - \lambda^T (h(x) - b) \quad (2.1)$$

where $\lambda \in R^m$ is a vector of **Lagrange multipliers**

THEOREM 2.3 (Lagrangian Sufficiency Theorem). *Let $x \in X$ and $\lambda \in R^m$ such that*

$$L(x, \lambda) = \inf_{x' \in X} L(x', \lambda) \quad (2.2)$$

and $h(x) = b$. Then x is optimal for P .

PROOF.

$$\min_{x' \in X(b)} f(x') = \min_{x' \in X(b)} [f(x') - \lambda^T (h(x') - b)] \quad (2.3)$$

$$\geq \min_{x' \in X} [f(x') - \lambda^T (h(x') - b)] \quad (2.4)$$

$$= f(x) - \lambda^T (h(x) - b) \quad (2.5)$$

$$= f(x) \quad (2.6)$$

□

2. Lagrange Dual

DEFINITION 2.4. Let

$$\phi(b) = \inf_{x \in X(b)} f(x). \quad (2.7)$$

Define the Lagrange dual function $g : R^m \rightarrow R$ with

$$g(\lambda) = \inf_{x \in X} L(x, \lambda) \quad (2.8)$$

Then, for all $\lambda \in R^m$,

$$\inf_{x \in X(b)} f(x) = \inf_{x \in X(b)} L(x, \lambda) \geq \inf_{x \in X} L(x, \lambda) = g(\lambda) \quad (2.9)$$

That is, $g(\lambda)$ is a lower bound on our optimization function.

This motivates the **dual problem** to maximize $g(\lambda)$ subject to $\lambda \in Y$, where $Y = \{\lambda \in R^m : g(\lambda) > -\infty\}$.

THEOREM 2.5 (Duality). *From (2.9), we see that the optimal value of the primal is always greater than the optimal value of the dual. This is **weak duality**.*

3. Supporting Hyperplanes

Fix $b \in R^m$ and consider ϕ as a function of $c \in R^m$. Further consider the hyperplane given by $\alpha : R^m \rightarrow R$ with

$$\alpha(c) = \beta - \lambda^T(b - c) \quad (2.10)$$

Now, try to find $\phi(b)$ as follow.

(i) For each λ , find

$$\beta_\lambda = \max\{\beta : \alpha(c) \leq \phi(c), \forall c \in R^m\} \quad (2.11)$$

(ii) Choose λ to maximize β_λ

DEFINITION 2.6. Call $\alpha : R^m \rightarrow R$ a **supporting hyperplane** to ϕ at b if

$$\alpha(c) = \phi(b) - \lambda^T(b - c) \quad (2.12)$$

and

$$\phi(c) \geq \phi(b) - \lambda^T(b - c) \quad (2.13)$$

for all $c \in R^m$.

THEOREM 2.7. *The following are equivalent*

- (i) *There exists a (non-vertical) supporting hyperplane to ϕ at b ,*
- (ii) *XXXXXXXXXXXXXXXXXXXXXXXXXXXX*

Linear Programming

1. Convexity and Strong Duality

DEFINITION 3.1. Let $S \subseteq R^n$. S is convex if for all $\delta \in [0, 1]$, $x, y \in S$ implies that $\delta x + (1 - \delta)y \in S$.

$f : S \rightarrow R$ is convex if for all $x, y \in S$ and $\delta \in [0, 1]$, $\delta f(x) + (1 - \delta)f(y) \geq f(\delta x + (1 - \delta)y)$.

Visually, the area under the function is a convex set.

DEFINITION 3.2. $x \in S$ is an interior point of S if there exists $\epsilon > 0$ such $\{y : \|y - x\|_2 \leq \epsilon\} \subseteq S$.

$x \in S$ is an extreme point of S if for all $y, z \in S$ and $S \in (0, 1)$, $x = \delta y + (1 - \delta)z$ implies $x = y = z$.

THEOREM 3.3 (Supporting Hyperplane). *Suppose that our function ϕ is convex and b lies in the interior of the set of points where ϕ is finite. Then there is a (non-vertical) supporting hyperplane to ϕ at b .*

THEOREM 3.4. *Let $X(b) = \{x \in X : h(x) \leq b\}$, $\phi(b) = \inf_{x \in X(b)} f(x)$. Then ϕ is convex if X , f , and h are convex.*

PROOF. Let $b_1, b_2 \in R^m$ such that $\phi(b_1), \phi(b_2)$ are defined. Let $\delta \in [0, 1]$ and $b = \delta b_1 + (1 - \delta)b_2$. Consider $x_1 \in X(b_1)$, $x_2 \in X(b_2)$ and let $x = \delta x_1 + (1 - \delta)x_2$.

By convexity of Y , $x \in X$. By convexity of h ,

$$\begin{aligned} h(x) &= h(\delta x_1 + (1 - \delta)x_2) \\ &\leq \delta h(x_1) + (1 - \delta)h(x_2) \\ &\leq \delta b_1 + (1 - \delta)b_2 = b \end{aligned}$$

Thus $x \in X(b)$, and by convexity of f ,

$$\begin{aligned} \phi(b) &\leq f(x) \\ &= f(\delta x_1 + (1 - \delta)x_2) \\ &\leq \delta f(x_1) + (1 - \delta)f(x_2) \\ &\leq \delta \phi(b_1) + (1 - \delta)\phi(b_2) \end{aligned}$$

□

2. Linear Programs

DEFINITION 3.5. General form of a linear program is

$$\min\{c^T x : Ax \geq b, x \geq 0\} \quad (3.1)$$

Standard form of a linear program is

$$\min\{c^T x : Ax = b, x \geq 0\} \quad (3.2)$$

3. Linear Program Duality

Introduce slack variables to turn problem \mathbb{P} into the form

$$\min\{c^T x : Ax - z = b, x, z \geq 0\} \quad (3.3)$$

We have $X = \{(x, z) : x, z \geq 0\} \subseteq R^{m+n}$. The Lagrangian is

$$L((x, z), \lambda) = c^T x - \lambda^T (Ax - z - b) \quad (3.4)$$

$$= (c^T - \lambda^T A)x + \lambda^T z + \lambda^T b \quad (3.5)$$

and has a finite minimum if and only if

$$\lambda \in Y = \{\lambda : c^T - \lambda^T A \geq 0, \lambda \geq 0\} \quad (3.6)$$

For a fixed $\lambda \in Y$, the minimum of L is satisfied when $(c^T - \lambda^T A)x = 0$ and $\lambda^T z = 0$, and thus

$$g(\lambda) = \inf_{(x, z) \in X} L((x, z), \lambda) = \lambda^T b \quad (3.7)$$

We obtain that the dual problem

$$\max\{\lambda^T b : A^T \lambda \leq c, \lambda \geq 0\} \quad (3.8)$$

and it can be shown (exercise) that the dual of the dual of a linear program is the original program.

4. Complementary Slackness

THEOREM 3.6. *Let x and λ be feasible for \mathbb{P} and its dual. Then x and λ are optimal if and only if*

$$(c^T - \lambda^T A)x = 0 \quad (3.9)$$

$$\lambda^T (Ax - b) = 0 \quad (3.10)$$

PROOF. If x, λ are optimal, then

$$c^T x = \underbrace{\lambda^T b}_{\text{by strong duality}} \quad (3.11)$$

$$= \inf_{x' \in X} (c^T x' - \lambda^T (Ax' - b)) \leq c^T x - \underbrace{\lambda^T (Ax - b)}_{\text{primal and dual optimality}} \leq c^T x \quad (3.12)$$

Then the inequalities must be equalities. Thus

$$\lambda^T b = c^T x - \lambda^T (Ax - b) = \underbrace{(c^T - \lambda^T A)x}_{=0} + \lambda^T b \quad (3.13)$$

and

$$c^T x - \underbrace{\lambda^T (Ax - b)}_{=0} = c^T x \quad (3.14)$$

If $(c^T - \lambda^T A)x = 0$ and $\lambda^T (Ax - b) = 0$, then

$$c^T x = c^T - \lambda^T (Ax - b) = (c^T - \lambda^T A)x + \lambda^T b = \lambda^T b \quad (3.15)$$

and so by weak duality, x and λ are optimal. \square

CHAPTER 4

Simplex Method

1. Basic Solutions

Maximize $c^T x$ subject to $Ax = b$, $x \geq 0$, $A \in \mathbb{R}^{m \times n}$.

Call a solution $x \in \mathbb{R}^n$ of $Ax = b$ **basic** if it has at most m non-zero entries, that is, there exists $B \subseteq \{1, \dots, n\}$ with $|B| = m$ and $x_i = 0$ if $i \notin B$.

A basic solution x with $x \geq 0$ is called a **basic feasible solution** (BFS).

2. Extreme Points and Optimal Solutions

We make the following assumptions:

- (i) The rows of A are linearly independent
- (ii) Every set of m columns of A are linearly independent.
- (iii) Every basic solution is non-degenerate - that is, it has exactly m non-zero entries.

THEOREM 4.1. *x is a BFS of $Ax = b$ if and only if it is an extreme point of the set $X(b) = \{x : Ax = b, x \geq 0\}$.*

THEOREM 4.2. *If the problem has a finite optimum (feasible and bounded), then it has an optimal solution that is a BFS.*

3. The Simplex Tableau

Let $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$. Let B be a basis (in the BFS sense), and $x \in \mathbb{R}^n$, such that $Ax = b$. Then

$$A_B x_B + A_N x_N = b \tag{4.1}$$

where $A_B \in \mathbb{R}^{m \times m}$ and $A_N \in \mathbb{R}^{m \times (n-m)}$ respectively consist of the columns of A indexed by B and those not indexed by B . Moreover, if x is a basic solution, then there is a basis B such that $x_N = 0$ and $A_B x_B = b$, and if x is a basic feasible solution, there is a basis B such that $x_N = 0$, $A_B x_B = b$, and $x_B \geq 0$.

For every x with $Ax = b$ and every basis B , we have

$$x_B = A_B^{-1}(b - A_N x_N) \tag{4.2}$$

as we assume that A_B has full rank. Thus,

$$f(x) = c^T x = c_B^T x_B + c_N^T x_N \quad (4.3)$$

$$= c_B^T A_B^{-1}(b - A_N x_N) + c_N^T x_N \quad (4.4)$$

$$= C_B^T A_B^{-1} b + (c_N^T - c_B^T A_B^{-1} A_N) x_N \quad (4.5)$$

Assume we can guarantee that $A_B^{-1} b \geq 0$. Then x^* with $x_B^* = A_B^{-1} b$ and $x_N^* = 0$ is a BFS with

$$f(x^*) = C_B^T A_B^{-1} b \quad (4.6)$$

Assume that we are maximizing $c^T x$. There are two different cases:

- (i) If $C_N^T - C_B^T A_B^{-1} A_N \leq 0$, then $f(x) \leq C_B^T A_B^{-1} b$ for every feasible x , so x^* is optimal.
- (ii) If $(C_N^T - C_B^T A_B^{-1} A_N)_i > 0$, then we can increase the objective value by increasing the corresponding row of $(x_N)_i$.

4. The Simplex Method in Tableau Form

Advanced Simplex Procedures

1. The Two-Phase Simplex Method

Finding an initial BFS is easy if the constraints have the form $Ax = b$ where $b \geq 0$, as

$$Ax + z = b, (x, z) = (0, b) \quad (5.1)$$

2. Gomory's Cutting Plane Method

Used in integer programming (IP). This is a linear program where in addition some of the variables are required to be integral.

Assume that for a given integer program we have found an optimal fractional solution x^* with basis B and let $a_{ij} = (A_B^{-1}A_j)$ and $a_{i0} = (A_B^{-1}b)$ be the entries of the final tableau. If x^* is not integral, then for some row i , a_{i0} is not integral. For every feasible solution x ,

$$x_i = \sum_{j \in \mathbb{N}} \lfloor a_{ij} \rfloor x_j \leq x_i + \sum_{j \in \mathbb{N}} a_{ij} x_j = a_{i0}. \quad (5.2)$$

If x is integral, then the left hand side is integral as well, and the inequality must still hold if the right hand side is rounded down.

Thus,

$$x_i + \sum_{j \in \mathbb{N}} \lfloor a_{ij} \rfloor x_j \leq \lfloor a_{i0} \rfloor. \quad (5.3)$$

Then, we can add this constraint to the problem and solve the augmented program. One can show that this procedure converges in a finite number of steps.

CHAPTER 6

Complexity of Problems and Algorithms

1. Asymptotic Complexity

We measure complexity as a function of input size. The input of a linear programming problem: $c \in R^n, A \in R^{m \times n}, b \in R^m$ is represented in $(n + m \cdot n + m) \cdot k$ bits if we represent each number using k bits.

For two functions $f : \mathbb{R} \rightarrow \mathbb{R}$ and $g : \mathbb{R} \rightarrow \mathbb{R}$ write

$$f(n) = \mathcal{O}(g(n)) \tag{6.1}$$

if there exists c, n_0 such that for all $n \geq n_0$,

$$f(n) \leq c \cdot g(n) \tag{6.2}$$

, ... (similarly for $\Omega \rightarrow \geq$, and $\Theta \rightarrow (\Omega + \mathcal{O})$)

CHAPTER 7

The Complexity of Linear Programming

1. A Lower Bound for the Simplex Method

THEOREM 7.1. *There exists a LP of size $\mathcal{O}(n^2)$, a pivoting rule, and an initial BFS such that the simplex method requires $2^n - 1$ iterations.*

PROOF. Consider the unit cube in \mathbb{R}^n , given by constraints $0 \leq x_i \leq 1$ for $i = 1, \dots, n$. Define a spanning path inductively as follows. In dimension 1, go from $x_1 = 0$ to $x_1 = 1$. In dimension k , set $x_k = 0$ and follow the path for dimension $1, \dots, k-1$. Then set $x_k = 1$, and follow the path for dimension $1, \dots, k-1$ backwards.

The objective x_n currently increases only once. Instead consider the perturbed unit cube given by the constraints $\epsilon \leq x_1 \leq 1$, $\epsilon x_{i-1} \leq x_i \leq 1 - \epsilon x_{i-1}$ with $\epsilon \in (0, \frac{1}{2})$. \square

2. The Idea for a New Method

$$\min\{c^T x : Ax = b, x \geq 0\} \tag{7.1}$$

$$\max\{b^T \lambda : A^T \lambda \leq c\} \tag{7.2}$$

By strong duality, each of these problems has a bounded optimal solution if and only if the following set of constraints is feasible:

$$c^T x = b^T \lambda \tag{7.3}$$

$$Ax = b \tag{7.4}$$

$$x \geq 0 \tag{7.5}$$

$$A^T \lambda \leq c \tag{7.6}$$

It is thus enough to decide, for a given $A' \in \mathbb{R}^{m \times n}$ and $b' \in \mathbb{R}^m$, whether $\{x \in \mathbb{R}^n : Ax \geq b\} \neq \emptyset$.

DEFINITION 7.2. A symmetric matrix $D \in \mathbb{R}^{n \times n}$ is called positive definite if $x^T D x > 0$ for every $x \in \mathbb{R}^n$. Alternatively, all eigenvalues of the matrix are strictly positive.

DEFINITION 7.3. A set $E \subseteq \mathbb{R}^n$ given by

$$E = E(z, D) = \{x \in \mathbb{R}^n : (x - z)^T D (x - z) \leq 1\} \tag{7.7}$$

for a positive definite symmetric $D \in \mathbb{R}^{n \times n}$ and $z \in \mathbb{R}^n$ is called an ellipsoid with center z .

Let $P = \{x \in \mathbb{R}^n : Ax \geq b\}$ for some $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$. To decide whether $P \neq \emptyset$, we generate a sequence $\{E_t\}$ of ellipsoids E_t with centers x_t . If $x_t \in P$, then P is non-empty and the method stops. If $x_t \notin P$, then one of the constraints is violated - so there exists a row j of A such that $a_j^T x_t < b_j$. Therefore, P is contained in the half-space $\{x \in \mathbb{R}^n : a_j^T x \geq a_j^T x_t\}$, and in particular the intersection of this half-space with E_t , which we will call a half-ellipsoid.

THEOREM 7.4. *Let $E = E(z, D)$ be an ellipsoid in \mathbb{R}^n and $a \in \mathbb{R}^n \neq 0$. Consider the half-space $H = \{x \in \mathbb{R}^n : a^T x \geq a^T z\}$, and let*

$$z' = z + \frac{1}{n+1} \frac{Da}{\sqrt{a^T D a}} \quad (7.8)$$

$$D' = \frac{n^2}{n^2 - 1} \left(D - \frac{2}{n+1} \frac{Daa^T D}{a^T D a} \right) \quad (7.9)$$

Then D' is symmetric and positive definite, and therefore $E' = E(z', D')$ is an ellipsoid. Moreover, $E \cap H \subseteq E'$ and $\text{Vol}(E') < e^{\frac{-1}{2(n+1)}} \text{Vol}(E)$.

CHAPTER 8

Graphs and Flows

1. Introduction

Consider a directed graph (network) $G = (V, E)$, V the set of vertices, $E \subseteq V \times V$ a set of edges. Undirected if E is symmetric.

2. Minimal Cost Flows

Fill in this stuff
from lectures

CHAPTER 9

Transportation and Assignment Problems

CHAPTER 10

Non-Cooperative Games

THEOREM 10.1 (von Neumann, 1928). *Let $P \in \mathbb{R}^{m \times n}$. Then*

$$\max_{x \in X} \min_{y \in Y} p(x, y) = \min_{y \in Y} \max_{x \in X} p(x, y) \quad (10.1)$$

CHAPTER 11

Strategic Equilibrium

DEFINITION 11.1. $x \in X$ is a best response to $y \in Y$ if $p(x, y) = \max_{x' \in X} p(x', y)$. $(x, y) \in X \times Y$ is a Nash equilibrium if x is a best response to y and y is a best response to x .

THEOREM 11.2. $(x, y) \in X \times Y$ is an equilibrium of the matrix game P if and only if

$$\min_{y' \in Y} p(x, y') = \max_{x' \in X} \min_{y' \in Y} p(x', y') \quad (11.1)$$

and

$$\max_{x \in X} p(x, y) = \min_{y' \in Y} \max_{x' \in X} p(x', y'). \quad (11.2)$$

THEOREM 11.3. Let $(x, y), (x', y') \in X \times Y$ be equilibria of the matrix game with payoff matrix P . Then $p(x, y) = p(x', y')$ and (x, y') and (x', y) are equilibria as well.

THEOREM 11.4 (Nash, 1951). Every bimatrix game has an equilibrium.

Bibliography