Optimization

Linear Programming

$$\begin{array}{llll} \text{P: maximize } c^T x & \text{s.t.} & Ax \le b, & x \ge 0 \\ \text{D: minimize } \lambda^T b & \text{s.t.} & A^T\lambda \le c, & \lambda \ge 0 \end{array}$$

The simplex algorithm. Slack variables. The two-phase algorithm — artificial variables. Shadow prices.

Complementary slackness

P		D
variables x		constraints λ
x_i basic ($x_i \neq 0$)	$\Longrightarrow$	constraint: tight ($v_i = 0$)
x_i non-basic ($x_i = 0$)	$\Longleftarrow$	constraint: slack ($v_i \neq 0$)
constraints		variables λ
constraint: tight ($z_i = 0$)	$\Longleftarrow$	λ_i basic ($\lambda_i \neq 0$)
constraint: slack ($z_i \neq 0$)	$\Longrightarrow$	λ_i non-basic ($\lambda_i = 0$)

Theorem

The feasible set of an LP problem is convex.

Proof

Write the problem as minimize $c^T x$ s.t. $Ax = b$, $x \ge 0$, and let X_b be the feasible set. Suppose $x, y \in X_b$, so $x, y \ge 0$ and $Ax = Ay = b$. Consider $z = \lambda x + (1-\lambda)y$ for $0 \le \lambda \le 1$. Then $z_i = \lambda x_i + (1-\lambda)y_i \ge 0$ for each i. So $z \ge 0$. Secondly,

$$Az = A(\lambda x + (1-\lambda)y) = \lambda Ax + (1-\lambda)Ay = \lambda b + (1-\lambda)b = b.$$

So $z \in X_b$ and hence X_b is convex.

Theorem

Basic feasible solutions $\equiv$ extreme points of the feasible set.

Proof

Suppose x is a b.f.s. Then $\exists$ a basis B and non-basis N such that the non-basic components of x satisfy $x_N = 0$. Suppose x is not extreme. Then $\exists y, z$ such that x lies on the line segment between y and z and hence $y_N = z_N = 0$ (proof omitted). Furthermore, y and z are feasible, so $A_B y_B = A_B z_B = b$ and this gives $A_B(y_B - z_B) = 0$. But as A_B is non-singular (by assumption) we have $y_B = z_B$ and hence $y = z$. This is a contradiction, and so x must be extreme.

Now suppose x is an extreme point of X_b. Since $x \in X_b$ we know it is feasible — we just need to show that it is basic.

Suppose it is not a b.f.s. Then the number of non-zero coordinates, p say, is greater than m. Let $P = \{i : x_i > 0\}$ and $Q = \{i : x_i = 0\}$. Since x is feasible, $Ax = A_P x_P + A_Q x_Q = b \Rightarrow A_P x_P = b$. But this is m equations in $p > m$ variables, so $\exists$ non-zero y_P s.t. $A_P y_P = 0$. We put $y_Q = 0$ and $y = \binom{y_P}{y_Q}$.

Consider $x + \epsilon y$ and $x - \epsilon y$ for small ϵ. For $\epsilon > 0$ small enough these two points are both feasible since $A(x \pm \epsilon y) = Ax \pm \epsilon Ay = b$, and $x \pm \epsilon y \geq 0$ (for small ϵ since $y_i > 0$ implies $x_i > 0$). Hence

$$x = \tfrac{1}{2}(x + \epsilon y) + \tfrac{1}{2}(x - \epsilon y)$$

and so x is not extreme. This is a contradiction, and hence x must be a b.f.s.

Theorem

If an LP has a finite optimum then there is an optimal basic feasible solution.

Proof

See lecture notes.

Theorem (weak duality)

If x is feasible for P and λ is feasible for D then $c^T x \leq b^T \lambda$. In particular, if one problem is feasible then the other is bounded.

Proof

Let $L(x, z, \lambda) = c^T x - \lambda^T (Ax + z - b)$ where $Ax + z = b$. Now for x and λ satisfying the conditions of the theorem,

$$c^T x = L(x, z, \lambda) = (c^T - \lambda^T A)x - \lambda^T z + \lambda^T b \leq \lambda^T b.$$

Theorem (sufficient conditions for optimality)

If x^*, z^* are feasible for P and λ^* is feasible for D, and x^*, z^*, λ^* satisfy complementary slackness, then x^* is optimal for P and λ^* is optimal for D. Furthermore $c^T x^* = \lambda^{*T} b$.

Proof

Let $L(x, z, \lambda) = c^T x - \lambda^T (Ax + z - b)$. Then

$$\begin{aligned} c^T x^* &= L(x^*, z^*, \lambda^*) \\ &= (c^T - \lambda^{*T} A)x^* - \lambda^{*T} A x^* + \lambda^{*T} b \\ &= \lambda^{*T} b \end{aligned}$$

by complementary slackness. But for all x feasible for P we have $c^T x \leq \lambda^{*T} b$ (by the weak duality theorem) and this implies that for all feasible x, $c^T x \leq c^T x^*$. So x^* is optimal for P. Similarly λ^* is optimal for D, and the problems have the same solutions.

Theorem (strong duality: necessary conditions for optimality)

If both P and D are feasible then $\exists\, x, \lambda$ satisfying the conditions above.

Lagrangian Methods

The Lagrangian

For the general optimization problem

$$\text{P: minimize } f(x) \quad \text{s.t.} \quad g(x) = b,\ x \in X$$

the Lagrangian is

$$L(x, \lambda) = f(x) - \lambda^T (g(x) - b).$$

The Lagrangian sufficiency theorem

If x^* and λ^* exist such that x^* is feasible for P and

$$L(x^*, \lambda^*) \leq L(x, \lambda^*) \quad \forall x \in X,$$

then X^* is optimal for P.

Proof

Define

$$X_b = \{x : x \in X \text{ and } g(x) = b\}.$$

Note that $X_b \subseteq X$ and that for any $x \in X_b$

$$L(x, \lambda) = f(x) - \lambda^T (g(x) - b) = f(x).$$

Now

$$f(x^*) = L(x^* \lambda^*) \leq L(x, \lambda^*) = f(x), \quad \forall x \in X_b.$$

Thus x^* is optimal for P.

Theorem (weak duality)

For $\lambda \in Y$, let

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

Then for any $x \in X_b$, $\lambda \in Y$,

$$L(\lambda) \leq f(x).$$

Proof

For $x \in X_b$, $\lambda \in Y$,

$$f(x) = L(x, \lambda) \geq \min_{x \in X_b} L(x, \lambda) \geq \min_{x \in X} L(\lambda, x) = L(\lambda).$$

Solution of general optimization problems

1. Find Y, the set of λ such that $L(x, \lambda)$ has a finite minimum.
2. Find $x(\lambda)$, the value of x at which this minimum is obtained, for all $\lambda \in Y$.
3. Find $x^* = x(\lambda^*)$, where x^* is feasible for P.
4. Then x^* is optimal for P.

Applications

Two person zero-sum games

Let A be the pay-off matrix for the game. A pair of strategies p and q with $\sum p_i = q_i = 1$, $p_i \geq 0$ and $q_i \geq 0$ are optimal if

$$p^T A \geq v, \qquad Aq \leq v \qquad \text{and} \qquad p^T Aq = v$$

for some $v \in \mathbb{R}$, where v is the value of the game.

The max flow/min cut theorem

The maximal flow value through a network is equal to the minimum cut capacity.

Proof

Define $f(X, Y) = \sum_{i \in X, j \in Y} x_{ij}$, the flow from X to Y. Let the set of nodes of the network be represented by $N = \{1, 2, \ldots, n\}$. Let $(S, \bar{S})$ be a cut and (x_{ij}) be a feasible flow with value v. Note that $f(X, N) = f(X, S) + f(X, \bar{S})$. Since (x_{ij}) is feasible, we have

$$\sum_j x_{ij} - \sum_j x_{ji} = \begin{cases} v & i = 1 \\ 0 & i \neq 1, n \\ -v & i = n \end{cases}$$

Summing this equality over $i \in S$ we obtain

$$\begin{aligned} v &= \sum_{i \in S} \sum_j (x_{ij} - xji) \\ &= \sum_{i \in S} \sum_j x_{ij} - \sum_{i \in S} \sum_j x_{ji} \\ &= f(S, N) - f(N, S) \\ &= f(S, S) + f(S, \bar{S}) - f(\bar{S}, S) - f(S, S) \\ &= f(S, \bar{S}) - f(\bar{S}, S) \\ &\leq f(S, \bar{S}) \\ &\leq C(S, \bar{S}) \end{aligned}$$

where $C(S, \bar{S})$ is the capacity of the cut $(S, \bar{S})$. So any flow is $\leq$ any cut capacity, and in particular the maximal flow is $\leq$ the minimal cut capacity.

Now let f be a maximal flow, and define $S \subseteq N$ recursively by:

1. $1 \in S$
2. $i \in S$ and $x_{ij} < c_{ij} \implies j \in S$
3. $i \in S$ and $x_{ji} > 0 \implies j \in S$

So S is the set of nodes to which we can increase the flow. Now if $n \in S$ then we can increase the flow along some path to n and so the flow is not maximal. Hence $n \in \bar{S} = N \setminus S$ and so

$(S, \bar{S})$ is a cut. From the definition of S we know that for $i \in S$ and $j \in \bar{S}$ we have $x_{ij} = c_{ij}$ and $x_{ji} = 0$, so in the formula above we get

$$v = f(S, \bar{S}) - f(\bar{S}, S) = f(S, \bar{S}) = C(S, \bar{S}).$$

Therefore the maximal flow is equal to the minimal cut capacity.

Sufficient conditions for a minimum cost circulation

If (x_{ij}) is a feasible circulation and there exists λ such that

$$x_{ij} = \begin{cases} c_{ij}^- & \text{if} \quad d_{ij} - \lambda_i + \lambda_j > 0 \\ c_{ij}^+ & \text{if} \quad d_{ij} - \lambda_i + \lambda_j < 0 \end{cases}$$
$$c_{ij}^- \leq x_{ij} \leq c_{ij}^+ \quad \text{if} \quad d_{ij} - \lambda_i + \lambda_j = 0$$

then (x_{ij}) is a minimal cost circulation. The λ_i are known as *node numbers* or *potentials*. In particular, if $c_{ij}^- = 0$ and $c_{ij}^+ = \infty$ then the conditions are

$$d_{ij} - \lambda_i + \lambda_j \geq 0$$
$$(d_{ij} - \lambda_i + \lambda_j)x_{ij} = 0.$$

Proof

Apply the Lagrangian sufficiency theorem.

The transportation algorithm

1. Pick an initial feasible solution with $m + n - 1$ non-zero flows (NW corner rule).
2. Set $\lambda_1 = 0$ and compute λ_i, μ_i using $d_{ij} - \lambda_i + \mu_j = 0$ on arcs with non-zero flows.
3. If $d_{ij} - \lambda_i + \mu_j \geq 0$ for all (i, j) then the flow is optimal.
4. If not, pick (i, j) for which $d_{ij} - \lambda_i + \mu_j < 0$.
5. Increase flow in arc (i, j) by as much as possible without making the flow in any arc negative. Return to 2.